Applying Functional Genomics to Understand the Virulence of Staphylococcus aureus

Yokoyama, Maho

Award date:
2018

Awarding institution:
University of Bath

Applying Functional Genomics to Understand the Virulence of *Staphylococcus aureus*

A thesis submitted for the degree of Doctor of Philosophy

Maho Yokoyama

Milner Centre for Evolution
Department of Biology and Biochemistry
University of Bath

May 2018

First of all, I would like to thank my supervisor Dr Ruth Massey – you have been brilliant, and your support has got me through this Ph.D. with my passion for research still intact!

Also, thanks to all the collaborators who have contributed to my work, especially Dr Maisem Laabei, Dr Sion Bayliss, Dr Harry Thorpe, Ms. Emily Stevens, Dr Leann Bacon, Dr Mario Recker, Dr Anneke Lubben, Dr Shaun Reeksting and Dr Ewan Murray.

"If you've a lesson to teach me I'm listening, ready to learn"

Alex Turner (Arctic Monkeys)

Finally, thanks to my family (especially Mum!) and friends for your support over the last four years – especially those that are or have been in the office/lab!

"You make me smile"

Miles Kane

Contents

Abstract

Staphylococcus aureus is asymptomatically carried in approximately 30% of the population, which is a risk for subsequent infections. *S. aureus* produces a vast array of virulence factors, leading to varied infections. In this project, a functional genomics approach was used to try and gain further insights into how this pathogen regulates its virulence. First, a result from a Genome Wide Association Study (GWAS) investigating the link between genotypic variation and variation in toxin production showed a link between mupirocin resistance and toxin production. Second, a group of clinical Methicillin-Resistant *S. aureus* (MRSA) isolates of the ST239 lineage were screened for variation in lipase activity, and a GWAS was carried out using this data.

Mupirocin is an antibiotic which targets isoleucyl-tRNA synthetase, which binds isoleucine to its tRNA to form isoleucyl-tRNA. The mutation associated with toxin production in the GWAS was shown to increase the presence of free isoleucine in mupirocin resistant *S. aureus* cells, suggesting that there was less isoleucyl-tRNA. This would be expected to affect proteins which are high in isoleucine, of which AgrC is one; this is part of the *accessory gene regulator* (*agr*) system, which controls virulence. Initially, AgrC is translated slower in mupirocin resistant strains, however there is no difference in overall *agr* activity at later time points. It was also found that competitive fitness of mupirocin resistant strains lacking the *agr* system is lower, therefore lowering toxin production seems to alleviate the fitness cost of this mutation.

The GWAS carried out on lipase production resulted in a list of genes, three of which were taken for further investigation - these were SAUSA300_1966, *murA* and *atpH*. SAUSA300_1966 is a putative phage-antirepressor protein, which may also repress lipase or interact with a lipase repressor. *murA* is part of peptidoglycan biosynthesis, and *atpH* is an ATP synthase subunit.

These studies demonstrate that a GWAS approach can be used to study the virulence of *S. aureus* by identifying potential regulators which can then be experimentally verified.

Chapter 1

- Introduction

Antibiotic resistance, *Staphylococcus aureus* carriage and infections

Antibiotic resistance in Staphylococcus aureus

Complicating the treatment of infections, *S. aureus* has become resistant to multiple antibiotics, with Methicillin-Resistant *S. aureus* (MRSA) being the most notable example[1,2]. This problem is exacerbated by inappropriate antibiotic use, ease and decreasing cost of international travel and the use of antibiotics as growth promoters for livestock. Antibiotic resistance can be conferred by various mechanisms, including mutation of the target so that the antibiotic no longer affects its action, the acquisition of resistance genes on plasmids or mobile genetic elements (MGEs) and by the use of efflux pumps to force the antibiotic out of the cell.

Carriage of S. aureus the risk of subsequent infection and atopic dermatitis

- *Risk of infections due to* S. aureus *carriage*

Staphylococcus aureus is a Gram-positive opportunistic pathogen most commonly found on the anterior nares of humans[3]. It can be asymptomatically carried, but also cause infections[4]. Carriage of *S. aureus* has been known to be a risk factor for *S. aureus* infections since 1931[5], and is thought to be the first step in the infection process[6]. People who are carriers have a higher rate of *S. aureus* infections than non-carriers, due mostly to their own strains[7], and breakdown of the epithelial barrier predisposes to intravascular/systemic infections[8]. Consequently, there does not seem to be a genetic difference between carriage isolates and those which cause infections, and there has been a case where the carriage isolate has gone on to cause bacteraemia with only a small number of mutations[9]. Temporary eradication of *S. aureus* from the nose has been demonstrated to reduce the rate of *S. aureus* infection in carriers[10].

- *Types of carriers*

There are three different carriage states; persistent, intermittent and non-carriage[11,12]. Depending on the method used, the percentage of the population falling into each of the carriage types varies, but generally around 20% are persistent nasal carriers, around 30% are intermittent carriers and around 50% are non-carriers[5]. Persistent carriers are colonised by one strain, while intermittent carriers may carry different strains at different times[13–15]. In one study, where a group of volunteers were inoculated with a *S. aureus* mixture, it was found that persistent carriers selectively re-acquire their strain while non-carriers quickly eliminate the mixture, suggesting involvement of host factors[11]. This is supported by the observation that there is variation in carriage rates between ethnicity[16], gender[16] and co-morbidities, etc. Factors which are associated with an increase in carriage rate include diabetes[17], IV drug use[18], and immune deficiency[18] – this includes Acquired Immune Deficiency Syndrome (AIDS), and quantitative or qualitative leukocyte deficiency. Another factor which seems to differentiate between persistent carriers and intermittent carriers is the load; persistent carriers have a higher *S. aureus* load compared to intermittent carriers[19].

- *How does* S. aureus *colonise the host?*

To successfully colonise the host, *S. aureus* needs to adhere to the epithelial cells in the nose. These are fully keratinised and the layers in the nose formed by these epithelial cells include several glands (apocrine sweat glands, sebaceous glands) and hair follicles[5]. The top-most layer is made up of keratinocytes, which are cornified – the cytoplasmic membrane is replaced by a cornified envelope – and these cells are eventually desquamated from the epithelium[20]. Loricrin accounts for the majority of the cornified envelope protein, making up around 85%, which cross-links with other proteins including keratins (cytokeratin 10, involucrin, filaggrin) and small proline-rich proteins, and these create a waterproof layer on the cornified envelope[20]. Keratinocytes have been shown to be important sites for *S. aureus* binding, and several *S. aureus* adhesins are able to bind to the cornified envelope; they include iron-regulated surface determinant A (IsdA) and clumping factor B (ClfB)[21–23]. Studies have shown ClfB to be the major adhesion in *S. aureus* attachment to the cornified envelope, interacting with cytokeratin 10 and loricrin[21]. The expression of these cornified envelope proteins is linked to the tissue repair process, and controlled by cytokines including ilterleukin-6 (IL-6) and IL-

22[24]; IL-6 production is triggered by the interaction between keratinocytes and *S. aureus*, and when heat-killed and sonicated *S. aureus* cells interact with these cells IL-6 production leads to a reduction in loricrin, cytokeratin 10 and filaggrin[25]. IL-22 meditates keratinocyte differentiation, and this leads to a decreased expression of the above cornified envelope proteins[26]. Therefore, a reduction in IL-6 and IL-22 is required in order for the cornifies envelope proteins to be expressed, which then allows for adhesion of *S. aureus* to establish colonisation[24,27].

In order to establish carriage, *S. aureus* must interact with the host immune system in such a way that it is not cleared. Nasal secretion, or mucous, contains antimicrobial factors, including IgA, IgG, lysozyme, lactoferrin and antimicrobial peptides[5]. Some antimicrobial peptides have been shown to be inactive towards *S. aureus*, or require the presence of other peptides to have activity towards *S. aureus*[5,28,29]. *S. aureus* commonly colonises sites in the nares where there are reduced nasal cilia and where there is less mucous[16]. It has been noted that *S. aureus* carriers have higher concentrations of α-defensins Human Neutrophil Peptide (HNP) 1, HNP2 and HNP3, and human β-defensin 2 (HBD2), which suggests that neutrophil- and epithelial-mediated inflammation is active[30]. Knowing that lipoteichoic acid stimulates neutrophil chemotaxis[31], the indication is that *S. aureus* colonisation causes an immune reaction, however, the HNP1, -2, -3 and HBD2 do not kill *S. aureus in vitro*. Therefore whatever immune response is mediated by the carriers is not effective at eradicating *S. aureus*[16]. Putting these together, the innate immune response against *S. aureus* in the nares is not efficient, but may play a role in ensuring that the carriage strain does not cause in an invasive infection[5].

- *Does carriage of* S. aureus *influence immunity?*

Differences in adaptive immunity between carriers and non-carriers have also been noted; carriers have been found to have a higher titre of IgG in their serum compared to non-carriers[12,32], and persistent carriers have a higher concentration of IgA specific for staphylococcal proteins[33]. Despite this, antibody responses between individuals (carriers and non-carriers) display significant variation[24]. This antibody response may provide some protection against *S. aureus* infections in carriers[34], whose antibodies are able to neutralise superantigens produced by their carriage strain[35]. Carriers also display a higher IgG titre after

infection with their carriage strain – these show similar binding pattern before and after infection[36]. While this may suggest that carriers may be "pre-immunised" with their carriage strain[36], when human volunteers were colonised short-term with *S. aureus* there was no evidence of an antibody response[37]. This shows that while carriage does cause a change in the humoral immune response, this may not be protective and also this may not influence carriage; maternal IgG does not seem to protect infants from *S. aureus* carriage[38].

Cellular adaptive immunity has been shown to be required for *S. aureus* clearance from the nose, with the Th17 cytokines IL-17 and IL-22 being shown to promote clearance in several studies in animals[24,27]. IL-22 was also shown to induce the innate immune response by up-regulating the expression of antimicrobial peptides, as well as controlling the expression of loricrin and cytokeratin 10 as mentioned previously[27]. Interestingly, in humans those who had been able to clear *S. aureus* showed an elevated IFN-γ response, indicative of a Th1 response, therefore a Th1 response may be crucial in humans for driving *S. aureus* clearance[39].

- *Influence of the microbiome on* S. aureus *carriage*

As there are other bacteria colonising humans, studies have investigated how this affects *S. aureus* colonisation; these bacteria include *Corynebacterium* species, *Propionibacterium* species, coagulase-negative Staphylococci (CoNS, mainly *S. epidermidis*), and *Streptococcus pneunomiae*. There seems to be an inverse correlation between carriage of *S. aureus* and other commensal bacteria, as demonstrated by several studies[40–42]. Perhaps this is due to the fact that these bacteria need to compete to colonise the niche; the increase in the use of streptococcal vaccines, which prevents *S. pneumoniae* carriage, has led to increased *S. aureus* carriage in the vaccinated population[43]. Another study has found that H_2O_2 produced by *S. pneumoniae* has the ability to induce lysogenic phages in *S. aureus*, a mechanism by which *S. pneumoniae* eliminates *S. aureus* from the niche[44]. One study by Ramsey *et al*. found that in the presence of *Corynebacterium striatum*, *S. aureus* gene expression was altered; virulence genes were down-regulated while genes involved in colonisation were up-regulated[45]. This included a strong repression of the *accessory gene regulator* (*agr*) response[45] – the system will be explored further on. This led to increased adherence to epithelial cells and a decrease in haemolysin production[45], which is indicative of a shift towards commensalism. During *in vivo*

co-infection with *C. striatum*, *S. aureus* had lower fitness compared to a mono-infection, which supports the above finding[45].

- *Carriage of* S. aureus *and atopic dermatitis*

Atopic dermatitis is a chronic inflammatory skin disease found in people whose skin barrier is disrupted[46]. It is characterised by a polarised T-helper 2 (Th2) inflammation[47], typically characterised by eczematous lesions, itching and increased susceptibility to bacterial skin infections, particularly by *S. aureus*[47]. Patients with atopic dermatitis have been known to be carriers of *S. aureus*, notably in the eczematous lesions, where it is present in up to 100% of patients and has a higher CFU count compared to non-lesional skin[48]; this perhaps explains the high incidences of *S. aureus* skin infections seen in patients with atopic dermatitis.

High levels of *S. aureus* colonisation and the subsequent reduction in skin microbiome diversity have been linked to atopic dermatitis flares[49]. However, it has been noted that treating the inflammation by the use of agents such as corticosteroids or calcineurin inhibitors reduce the amount of *S. aureus* on the atopic dermatitis lesions; therefore, it seems that inflammation itself is driving the proliferation of *S. aureus* in these lesions[48,50]. The inflammation leads to damage of the skin, thus exposing extracellular matrix proteins, giving a surface on which *S. aureus* can attach, and this is exacerbated by scratching of the skin[48,51,52]. This causes further damage and also promotes the release of cytokines, which upregulates the expression of extracellular matrix proteins[52]. This increased *S. aureus* adhesion did not occur in IL-4 knock-out mice and incubating the skin with IL-4 lead to increased *S. aureus* adhesion, therefore it is likely that a Th2 response drives the production of extracellular matrix proteins (particularly fibronectin) to promote increased adhesion[48]. Another study showed evidence of biofilms on the skin of patients with atopic dermatitis, and also that the protease staphopain B was present; staphopain B was shown to degrade the antimicrobial peptide LL-37 into shorter fragments, perhaps protecting the *S. aureus* from its action[46].

As *S. aureus* colonisation is linked to atopic dermatitis flares, studies have investigated the effects of decolonisation on disease severity. The use of antibiotics long-term is a concern for inducing resistance, as well as having a harmful effect on the other bacteria which form the skin microbiome[47] – these bacteria may in fact prevent aberrant *S. aureus* growth, so removal

of these bacteria may in fact exacerbate the situation. Therefore, recently a study began looking at the efficacy of a bacteriophage endolysin for the treatment of atopic dermatitis; Staphefekt SA.100 is a chimeric endolysin engineered to lyse staphylococcal cell membrane[53]. If this proves safe and efficacious, then it would not only provide an alternative treatment for atopic dermatitis but possibly *S. aureus* SSTIs.

Resistance to β-lactam antibiotics; the Rise and Fall of MRSA Clones

S. aureus was generally susceptible to most antibiotics, however the introduction of these agents into clinical use has fuelled the rise of resistance in *S. aureus*. The development of resistance to penicillin and methicillin in *S. aureus* can be seen in "waves"; the initial rise of penicillin-resistant strains after the introduction of penicillin into the clinic, then the second wave of Methicillin-Resistant *S. aureus* emerging in the 1960s after its introduction into the clinic[2]. These "archaic MRSA" strains were prevalent in European hospitals (Healthcare-Associated MRSA (HA-MRSA)), however they did not spread into the community and were seen sporadically in the US[2,54,55]. By the 1980s, these archaic MRSA clones were replaced by their descendants or other MRSA lineages which were more successful[56]. This third wave then spread to the US and the rest of the world by the mid 1980's, but this third-wave MRSA pandemic was still confined to hospitals and other healthcare institutions[2].

To treat this rising MRSA pandemic, the use of vancomycin increased as this was one of the only antibiotics which was still effective against MRSA[2]. However, this emergence of Vancomycin-Intermediate *S. aureus* (VISA, Minimum Inhibitory Concentration (MIC) 8 - 16µg/ml) and Vancomycin-Resistant *S. aureus* (VRSA, MIC ≥128µg/ml)[57,58]. Around the same time, the fourth wave began when MRSA started to spread in the community (Community-Associated MRSA (CA-MRSA)). This was first noticed in Western Australia in the 1990s, and between 1997 and 1999 there were cases of MRSA seen in children with no healthcare contact in the US[2,59,60]. Both these incidences were caused by MRSA strains which were unrelated to the HA-MRSA circulating around the time, therefore it would seem that methicillin resistance had jumped into Methicillin-Susceptible *S. aureus* (MSSA) strains which were circulating in the community or very distantly related to the HA-MRSA clones[2].

In the US, the CA-MRSA wave is in fact mediated by two different lineages; the early cases, such as those mentioned above, were caused by the USA400 lineage, and the emergence of USA300 in between 1999-2001 effectively replaced the USA400 lineage[2]. These CA-MRSA strains are noted for the increased virulence, and it has been noted that they produce many virulence factors at a high level. PVL is thought to be one of the contributing factors behind the increased virulence. At sub-lytic concentrations, PVL is thought to "prime" neutrophils by partially activating them, and this leads to the release of pro-inflammatory mediators such as leukotriene B4 and IL-8[61]. Another non-lytic activity of PVL is the induction of reactive oxygen species production from neutrophils when stimulated by *N*-formyl-methionly-leucyl-phenylalanine (fMLP). The USA300 clone is now the most common cause of SSTIs in the US, and the CA-MRSA clones have now started to be seen in the healthcare system[62,63].

Methicillin resistance is conferred by the *mecA* gene carried on Staphylococcal cassette chromosome *mec* (SCC*mec*), a mobile genetic element[1,2]. *mecA* encodes an alternative PBP2, PBP2a, which is not susceptible to β-lactam antibiotics. To date, 8 allotypes of SCC*mec* have been identified, designated SCC*mec* I – VIII. SCC*mec* II is larger, and associated with HA-MRSA strains. Due to its size, resistance to other antibiotics can be carried alongside the *mecA* gene in SCC*mec*II. CA-MRSA frequently carry the smaller SCC*mec* IV, and are generally susceptible to non-β-lactam antibiotics, however, resistance to other classes of antibiotics in CA-MRSA is also increasing[2,64]. A study in our lab has found that the expression of *mecA* resulted in a reduction in the expression of cytotoxins by preventing AIP from being detected by the *agr* system, which subsequently lead to reduced virulence in a mouse sepsis model[65]. The study also showed that while MecA expression from SCC*mec* II was higher than the expression from SCC*mec* IV, therefore the presence of SCC*mec* IV may allow for the maintenance of high virulence, which correlates with the increased virulence seen in CA-MRSA strains which frequently carry SCC*mec* IV. This could represent adaptations to the two environments, where high MecA expression leading to increased resistance to β-lactams is beneficial in the hospital environment, but this consequently drives down the expression of virulence factors which inhibits these strains from getting into the community.

- *Does MRSA carriage differ from MSSA carriage?*

One study looked at the carriage rates for MRSA and methicillin-sensitive *S. aureus* (MSSA) in a Taiwanese high school – the study found that of the 323 *S. aureus* strains isolated in the study (which represented 26.2% of nasal screening swabs), 278 were MSSA and 45 were MRSA[66]. This represents a carriage rate of 22.6% for MSSA and 3.7% for MRSA. A review of 31 studies showed that healthcare workers had an MRSA carriage rate of 1.8%, with nurses having the highest carriage rate (6.9%)[67]. Another review of 33 studies found that people with livestock contact had a higher rate of MRSA carriage (14.2%) compared to the general population, which ranged from 0.8 - 1.3%[68]. A further review showed that contact with healthcare or nursing homes and coming into contact with healthcare associated pathogens such as *Clostridium difficile* also increased the likelihood of MRSA carriage at time of hospital admission[69]. The same review also showed that congestive heart failure, diabetes, pulmonary disease, immunosuppression and renal failure were also associated with an increased likelihood of MRSA carriage at time of hospital admission. These studies show that MRSA carriage rate is lower than MSSA carriage rates, but that certain people are more likely to carry MRSA, such as farmers and patients with certain co-morbidities.

Reducing the risk of carriage; decolonisation and the role of resistance in decolonisation failure

To reduce the risk posed by the carriage of *S. aureus*, particularly Methicillin-Resistant *S. aureus* (MRSA), decolonisation strategies are in place to eradicate the colonising strain. These regimes typically involve the use of topical agents such as chlorhexidine and mupirocin.

- *Chlorhexidine*

Chlorhexidine is an antiseptic typically in the form of chlorhexidine gluconate. It is a water soluble cationic biguanide which binds to and disrupts the cell wall, altering the osmotic balance[70]. Chlorhexidine, which has a positive charge, is thought to be attracted to the negatively charged bacterial cell wall, where it then binds to the phospholipids in the cytoplasmic membrane[71]. This then increases the permeability of the cytoplasmic membrane, allowing small molecules such as K+ ions to leak out[71].

Tolerance to this antiseptic in *S. aureus* is by the utilisation of efflux pumps coded by the plasmid bourn *qacA*/*B* as well as *smr*[72–74]. In one study, it was found that the presence of the *qac* genes was associated with ciprofloxacin and erythromycin resistance[74], and other studies have shown links between these genes and resistance to other antibiotics. The presence of *qac* genes and increased tolerance to chlorhexidine also has implications for decolonisation, with studies showing that tolerance to chlorhexidine leads to decolonisation failure[74].

- *Mupirocin*

Mupirocin is a polyketide antibiotic produced by *Pseudomonas fluorescens*, and targets isoleucyl-tRNA synthetase (IleRS); it inhibits IleRS by binding to the active site where isoleucine binds as well as binding to the ATP-binding pocket[75]. Resistance to mupirocin exists at two levels, low-level (Minimum Inhibitory Concentration (MIC) 8 - 256µg/ml) and high-level (MIC ≥512µg/ml). Low-level mupirocin resistance is mediated through single nucleotide polymorphisms (SNPs) in the chromosomal *ileS* gene[75], which will be discussed in more detail in chapters 3 and 4, and high-level mupirocin resistance is mediated through the acquisition of an alternative IleRS carried on a plasmid, *mupA*[75,76] or *mupB*[77]. Resistance to mupirocin at both levels has been associated with decolonisation failure, particularly high-level resistance[10].

The presence of *mupA* was discovered when a study showed that strains which were highly resistant to mupirocin showed two distinct activity peaks on a radiometric IleRS activity assay[78]. Initially, the highly resistant strains were grown with mupirocin, but then they were sub-cultured into medium without mupirocin prior to the extraction of IleRS at late log-phase, where the highly resistant strains showed two IleRS activity peaks. This would indicate that the *mupA* IleRS is active alongside the chromosomal IleRS during the log-phase. The *mupA* gene is carried on various plasmids in different strains, and in one instance it was shown that one such plasmid carrying *mupA* had conjugated from *Staphylococcus epidermidis* within a carrier[79]. These plasmids can also carry genes conferring resistance to other antibiotics, such as tetracycline resistance seen in *mupA* containing plasmids studied by Needham *et al*[80].
However, the correlation between the *mupA* gene and high-level mupirocin resistance is not straightforward; several incidences of *mupA* being present on the chromosome have been

reported, with the first being reported by Ramsey *et al*. in 1996[81]. In this study, the strain displayed low-level mupirocin resistance as opposed to the expected high-level resistance. This was also observed in a study by Fujimura *et al*., in which they studied an MRSA strain from Japan with low-level mupirocin resistance[82]. However, in the study by Udo *et al.* showed that the presence of *mupA* on the chromosome also lead to high-level mupirocin resistance[83]. Also, in a study by Driscoll *et al*. it was found that loss of high-level mupirocin resistance was seen in *mupA* positive isolates, where a single deletion caused a frame-shift which rendered the protein inactive[84]. The authors observed that high-level mupirocin resistance was frequently re-established, with the mutated allele reverting back to wild-type. Therefore, if a low-level mupirocin resistant strain is *mupA* positive, there is a possibility that this strain may become highly resistant – this has implications for decolonisation, as higher resistance to mupirocin would lead to a less effective clearance of MRSA.

The second alternative IleRS, *mupB*, was found by Seah *et al*., where they investigated a case of an MRSA strain showing high-level mupirocin resistance in the absence of *mupA*[77]. The authors found that the *mupB* gene has only 45.5% DNA sequence identity with the chromosomal *ileS* gene, and has more similarity with the *mupA* gene, with 65.5% DNA sequence being identical. At the protein level, the MupB protein has 25.4% identity with IleRS and 58.1% identity with MupA, but despite this low homology the MupB protein contains conserved motifs found in class I tRNA synthetases.

- *Is resistance to these agents linked with decolonisation failure?*

Due to the fact that mupirocin is typically used as an ointment (Bactroban® (GlaxoSmithKline) is 20mg/*g*[85]), the application site has a high concentration of the antibiotic[86]. Therefore, low-level mupirocin resistance would not be expected to impede MRSA decolonisation. Indeed, several studies have questioned the clinical relevance of low-level mupirocin resistance[86–88]. Even so, a trend towards decolonisation failure were seen in multiple studies - for example, five studies looking at MRSA decolonisation showed that 24% (total 84) of patients with high-level mupirocin resistant MRSA achieved decolonisation, 29% (totla 103) of patients with low-level mupR MRSA achieved decolonisation compared to 62% (total 627) of patients with mupirocin susceptible MRSA[89]. Therefore, mupirocin resistance at low and high levels can lead

to failure[90], and if resistance to the two commonly used agents are linked, then this would make decolonisation failure more likely. Indeed, one study found that the combination of low-level mupirocin resistance alongside resistance to chlorhexidine was associated with MRSA decolonisation failure[91]. This in turn impacts on infection rates, with one study showing that patients who were successfully decolonised (104/268) of MRSA had no subsequent infection, while in 4.3% of those whose decolonisation failed (7/164) there was subsequent MRSA infection[92].

Vancomycin resistance

As mentioned above, the increase in the use of vancomycin to treat MRSA infections has led to a rise in VISA and VRSA. *S. aureus* modifies its cell wall to become resistant to vancomycin; one mechanism is thought to be by the modification of peptidoglycan biosynthesis, whereby VISA strains synthesise more peptidoglycan but with less cross-linking, leaving more D-Ala-D-Ala terminal peptide, the target for vancomycin, exposed[1,93]. This results in a thicker cell wall with irregular shape, and the increase in exposed D-Ala-D-Ala residues can then bind more vancomycin and trap it without affecting cell wall biosynthesis. The reduced cross-linking is thought to be the result of a reduction in L-glutamine residues, leading to reduced amidation of the L-glutamate in the pentapeptide bridge. VRSA is thought to be the result of *S. aureus* acquiring the *vanA* gene, probably from *Enterococcus faecalis*[1,94]. The terminal peptide is modified to D-Ala-D-Lac when exposed to low concentrations of vancomycin, and this makes the terminal peptide less susceptible to binding by vancomycin, therefore cell wall biosynthesis is not affected.

Resistance to fluoroquinolones

Fluoroquinolones were introduced in the 1980s, and *S. aureus* quickly acquired resistance to this class of antibiotics[1]. Fluoroquinolones target topoisomerase IV (which breaks concatenated DNA) or DNA gyrase (which relaxes supercoiled DNA), and mutations in these genes and the induction of an efflux pump have been associated with resistance to these antibiotics. The mutations can arise spontaneously, and exposure to fluoroquinolones at sub-therapeutic concentrations may induce these mutations in a sub-population of colonising *S. aureus*.[95] The presence of these resistant sub-populations and limited concentrations of

fluoroquinolones at the site of infection may lead to more resistance mutations accumulate, thus driving increased resistance.

***Types of infections caused by* S. aureus**

Despite the fact that there is a lot of research carried out, infections caused by staphylococci are still a significant problem[7]. These infections range from skin and soft tissue infections to endocarditis, bacteraemia, bone and joint infections, and pneumonia[6]. *S. aureus* is also capable of colonising prosthetic devices, such as endovascular catheters[96–98], subsequently leading to infections. Here, I will explore a few of these infections in more detail.

- *Skin and Soft Tissue Infections*

Community-Associated Methicillin-Resistant *Staphylococcus aureus* (CA-MRSA) strains are a major cause of skin and soft tissue infections (SSTIs)[64], but incidences of *S. aureus* SSTIs have been reported prior to the expansion of CA-MRSA clones[99]. Cutaneous abscesses are generally regarded as the hallmark of *S. aureus* SSTIs, however other manifestations have been seen[99–105]. In children, impetigo is generally caused by *S. aureus*, where it presents as bullous/papular legions which progress to crusted legions[99,106]. Usually, these infections are not accompanied by systemic symptoms. Other forms of SSTI include cellulitis, necrotising fasciitis, pyomyositis and surgical site infections[99]. A severe surgical site infection is mediastititis, a complication arising from median sternotomy, a cardiac surgical procedure[99,107].

A feature of *S. aureus* SSTIs, particularly furunculosis, or abscesses in the hair follicle, is recurrence, where patients have three or more incidences of furunculosis within a year[108]. Colonisation is thought to be a factor in the development of these recurrent furunculosis attacks, and as such decolonisation can be beneficial in preventing these recurrent attacks[108]. However, *S. aureus* can also survive on household surfaces[109], thus may re-colonise carriers who had been decolonised. As with other SSTIs, furunculosis is associated mainly with CA-MRSA strains, or strains carrying the Panton Valentine Leukocidin (PVL) [108,110]. Innate and T-cell mediated immune responses play an important role in protection from SSTIs[22], however it is possible that any antibody response mediated against *S. aureus* is not protective, due to

the immune evasion strategies utilised by *S. aureus*, such as Protein A[111], which will be discussed further on.

- S. aureus *bacteraemia*

In *S. aureus* bacteraemia, common origins of infection include vascular catheters (particularly those which are already infected), SSTIs, lung infections, bone and joint infections and infective endocarditis[99]. However, in ~25% of cases, however, there is no clear focus of infection[99]. While *S. aureus* infections can be an origin for subsequent bacteraemia, bacteraemia can also lead to *S. aureus* infections at secondary sites; this includes infective endocarditis, septic arthritis and osteomyelitis[112]. Further complicating this, *S. aureus* bacteraemia can result in sepsis, further complicating these infections[112,113]. Knowing that epidemiology of *S. aureus* bacteraemia is linked to other forms of *S. aureus* infection, changes in this epidemiology will likely result in a change in the epidemiology of *S. aureus* bacteraemia[99]. For example, improvements in the management of intravascular catheters have resulted in a reduction in catheter related infections, therefore this contributes to reduction in *S. aureus* bacteraemia originating from infected intravascular catheters[114].

One group who are at particular risk of bacteraemia are patients undergoing haemodialysis, and *S. aureus* is an established organism causing these catheter related bloodstream infections in these patients[99,115]. Most commonly, these bacteraemia are secondary to infection or inflammation around the catheter[116]. As a preventative measure, studies have investigated the efficacy of using antibiotics at the catheter exit site and found that use of agents such as mupirocin and polysporin were effective at reducing the incidences of catheter related bacteraemia[116]. However, with the prolonged use of antibiotics comes the risk of bacteria developing resistance to the agent used. Studies have also shown that, in the short term, coating the catheter with antibiotics or antiseptics was effective at reducing the incidences of catheter related bacteraemia[117,118], although it is unclear if this will also apply long-term[116].

- *Infective endocarditis*

S. aureus is the most common cause of infective endocarditis, with a study reporting that the proportion of infective endocarditis caused by *S. aureus* to be 32% in 2009, an increase from 24% in 1998[119]. Damage to the cardiac endothelium gives a niche for bacteria to colonise, and this can be caused by direct trauma or inflammation[99,120]. This exposes the subendothelium, and leads to production of thrombotic vegetations formed by the deposition of fibrin and platelets onto extracellular matrix proteins and tissue factors[99,120]. Cell wall associated factors on *S. aureus* allows the bacteria to attach to these vegetations, thus setting a focus of infection for infective endocarditis[99,121]. *S. aureus* is also the most common cause of prosthetic heart valve infections, shown to account for 23 – 33% of cases in these patients[122,123]. These infections are related to healthcare-associates bacteraemia, which is frequently caused by *S. aureus*; once the bacteria have gained access to the bloodstream, then it is possible for the prosthetic valve to be seeded with the bacteria, leading to infection[124]. A patient is most likely to develop prosthetic valve related infective endocarditis during the first year after the valve is implanted, likely due to ongoing contact with the healthcare system and incomplete endothelialisation of the new valve[123].

Pathogenicity of *S. aureus* – virulence factors

S. aureus has various factors for virulence, and these can be categorised into those involved in attachment, immune evasion and toxins (adhesion, evasion and toxicity)[6]. Depending on the function of the virulence factor, they can be expressed on the cell surface or secreted out of the cell[98]. Certain virulence factors have more than one role in causing disease, and multiple virulence factors can have the same role in virulence.

Adhesins

- *Microbial Surface Components Recognising Adhesive Matrix Molecules*

The majority of molecules involved in attachment are known as “Microbial Surface Components Recognising Adhesive Matrix Molecules” (MSRAMMs) and is essential for the attachment of *S. aureus* to host cells to initiate carriage or infection[51]. These MSCRAMMs bind to host proteins such as collagen, fibrinogen, and fibronectin, and have been shown to play an important role in invasive infections such as endovascular infections and septic

arthritis[125,126], as well as enabling *S. aureus* to adhere to the host proteins coating prosthetic devices[127,128]. Adhesion is the initial step in the formation of biofilms, where *S. aureus* produce an extracellular matrix and protect itself from the host defences and antibiotics[129]; this will be explored later on. This allows *S. aureus* to persist, making these kind of infections extremely difficult to treat.

MSCRAMMs are typically covalently anchored to the cell wall peptidoglycan, and contain a Sec-dependent secretory signal at the N-terminus and an LPXTG sortase cleavage site, hydrophobic domain and a positively-charged region at the C-terminus[51,130]. The hydrophobic domain and the positively-charged region maintains the protein in the cell membrane during secretion, and this allows the sortase to act on the protein[21]. MSCRAMMs are distinguished from other cell wall anchored proteins by the presence of IgG-folded domains, where at least two are present adjacent to each other, and this is where the ligand binding occurs[21].

Clumping Factors

Clumping factors A and B (ClfA, ClfB) are a prime example of an *S. aureus* MSCRAMM (fig. 1.1); these bind fibrinogen using dock, lock and latch and form clumps[21,131,132]. ClfA is expressed throughout *S. aureus* growth, while ClfB is only expressed in aerobically growing cells during early exponential phase[132,133]. It has been shown that ClfA and CLfB bind to different regions of fibrinogen, therefore the synergistic action of the two may promote stronger adhesion. The N-terminal A region of Clfs contain three N domains (N1, N2 and N3), with N2 and N3 having the characteristic IgG-folded domains[21]. The R region, which connects the A region with the wall-spanning region (W), in Clfs is formed of serine-aspartate repeats[21].

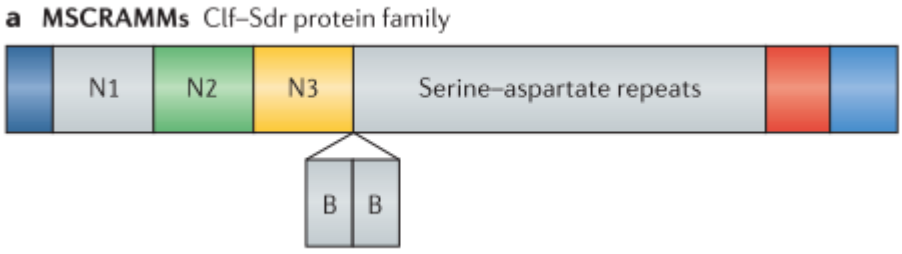


Fig. 1.1: a diagram showing the domains in clumping factors. Taken from Foster *et al*., Nat Rev Microbiol 2013[21].

Fibronectin Binding Proteins

Fibronectin binding proteins (FnBPs) are another member of MSCRAMMs (fig. 1.2). Two closely linked FnBPs are found in *S. aureus*, FnBPA and FnBPB[134], and they mediate attachment to immobilised fibronectin and plays a role in *S. aureus* adhesion to plasma clots and host conditioned implants – i.e. biomaterial which has resided in the host for a prolonged period, and thus coated with host extracellular matrix[97,135,136]. Like the Clfs, FnBPs have three N domains in their A region, but also contain Fibronectin-binding repeats in the R region[21]. As well as being a linker of A and wall-spanning regions, the R region of FnBPs forms the primary binding sites for fibronectin, however this domain seems to lack a secondary structure[51,137].

Fig. 1.2: a diagram showing the domains in fibronectin binding proteins. Taken from Foster *et al.*, Nat Rev Microbiol 2013[21].

Collagen binding protein

The collagen adhesin (Cna) is an MSCRAMM which binds to collagen, and is necessary for *S. aureus* to adhere to and cause cartilage infection (fig. 1.3). This MSCRAMM is not as frequently expressed as the other MSCRAMMs[51,138,139], and it would be assumed that this adhesin would be important in bone and joint infections; however, there is contradictory evidence for this, with one study suggesting that nearly all *S. aureus* isolates from these infections express Cna, but another study found no difference in the frequency of Cna expressing strains[51,138,139]. The A region of this protein is also divided into three domains, but unlike the other MSCRAMMs, the IgG-folded domains which bind to collagen is the N1 and N2 rather than the N2 and N3 used in Clfs and FnBPs[21]. The link between the A and W region comprises of a repeated B domains, and the linker between the B domains and the W domain is not flexible, differentiating Cna from the other MSCRAMMs[21].

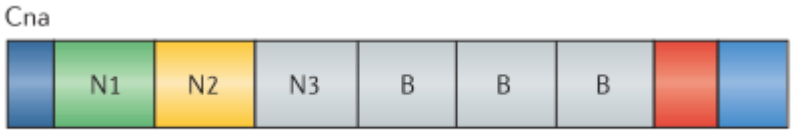


Fig. 1.3: a diagram showing the domains in the collagen adhesion protein Can. Taken from Foster *et al*., Nat Rev Microbiol 2013[21].

- *Other adhesins*

The NEAT motif family proteins include the iron-regulated surface proteins IsdA, IsdB and IsdH (fig. 1.4)[21]. These proteins contain at least one Near Iron Transporter (NEAT) motif, and these bind to haem which is then transported into the bacterial cell to release iron. This is an effective mechanism by which *S. aureus* acquires iron in an iron-limited environment, such as the bloodstream. IsdA has also been shown to enhance *S. aureus* attachment to squamous cells. Another group are the G5-E repeat family, which includes *S. aureus* surface protein G (SasG)[21]. These contain multiple G5-domains, which contain 5 conserved glycines and fold into a β-triple helix-β -like conformation. The G5 domains are separated by a 50-amino acid E regions. While SasG has no know ligand, it has been shown to be involved in adhesion to desquamated epithelial cells.

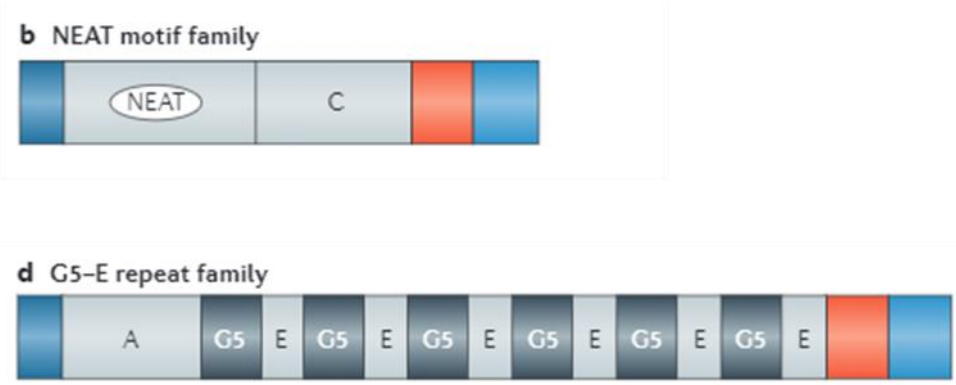


Fig. 1.4: a diagram showing the domains in a NEAT motif family protein and G5-E repeat family proteins. Taken from Foster *et al*., Nat Rev Microbiol 2013[21].

Immune Evasion

There are various mechanisms used by *S. aureus* to evade the immune system; *S. aureus* is capable of lysing immune cells by producing toxins (which will be explored further on),

inactivating complement, preventing the recruitment (chemotaxis) of neutrophils, resistance to phagocytosis and preventing the action of antimicrobial peptides[140]. *S.* aureus also produces factors including Protein A and capsule, as well as other cell wall-anchored proteins mentioned above to evade the immune system[140], and biofilm formation provides protection by the formation of an extracellular matrix[129].

- *Complement, complement evasion and chemotaxis inhibition*

The complement pathways are a part of the immune response in the host, and can be activated by the innate and adaptive immune response[141]. The main function of the pathway is the formation of the C3 convertase, leading to C3a, C3b, C5a and C5b formation. C3a is also known as anaphylatoxin and along with C3b, which coats, or opsonises, the target cell, causes the recruitment of immune cells as well as inducing a pro-inflammatory response. These C3b incorporates into the C3 convertase, leading to an increase in C3 convertases. This leads to the formation of the membrane attack complex (MAC) (fig. 1.5); C5b binds to C6 and C7, then deposits into the cell membrane. This leads to C8 binding to the C5b-C6-C7 complex, which then causes the recruitment and binding of C9 to form the lytic MAC.

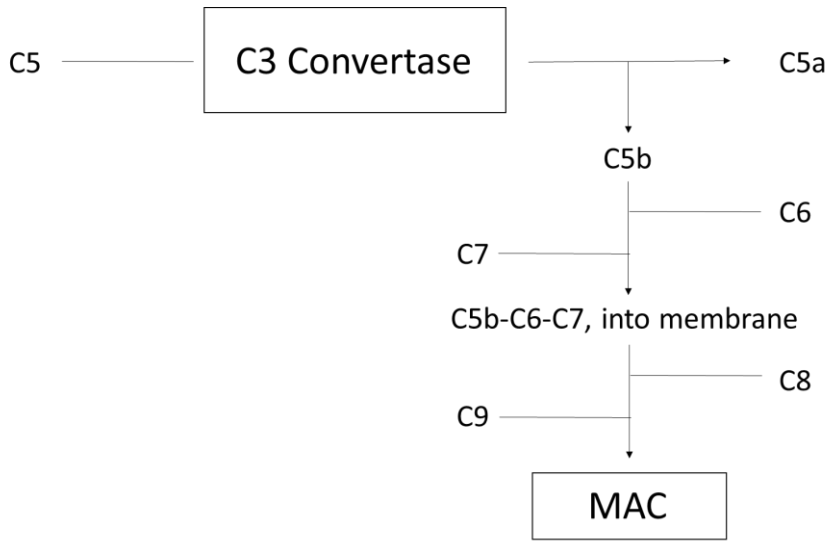


Fig. 1.5: a schematic showing the pathway leading to cell lysis by the complement pathways.

The classical pathway and the lectin pathway are related, producing the same "classical C3 convertase", C4b2b, but is activated by separate pathways (fig. 1.6)[141]. C1q is part of the C1 complex, and is the molecule in the classical pathway that recognises and binds to antibody

complexes, which activates the proteases C1r and C1s. This activated C1 complex then cleaves C4 into C4a and C4b, and exposes a thioester within C4b and thus C4b deposits on the target cell. The active C1 also cleaves C2, which is bound to C4b, into C2a and C2b, and this then becomes the C4b2b C3 convertase. In the lectin pathway, mannose binding lectin (MBL) and ficolin, which recognises carbohydrates, recruits MBL associated serine proteases (MASPs), which shares function with C1r and C1s. MASP-2 is the only MASP which cleaves C4 and C2, which gives rise to the classical C3 convertase.

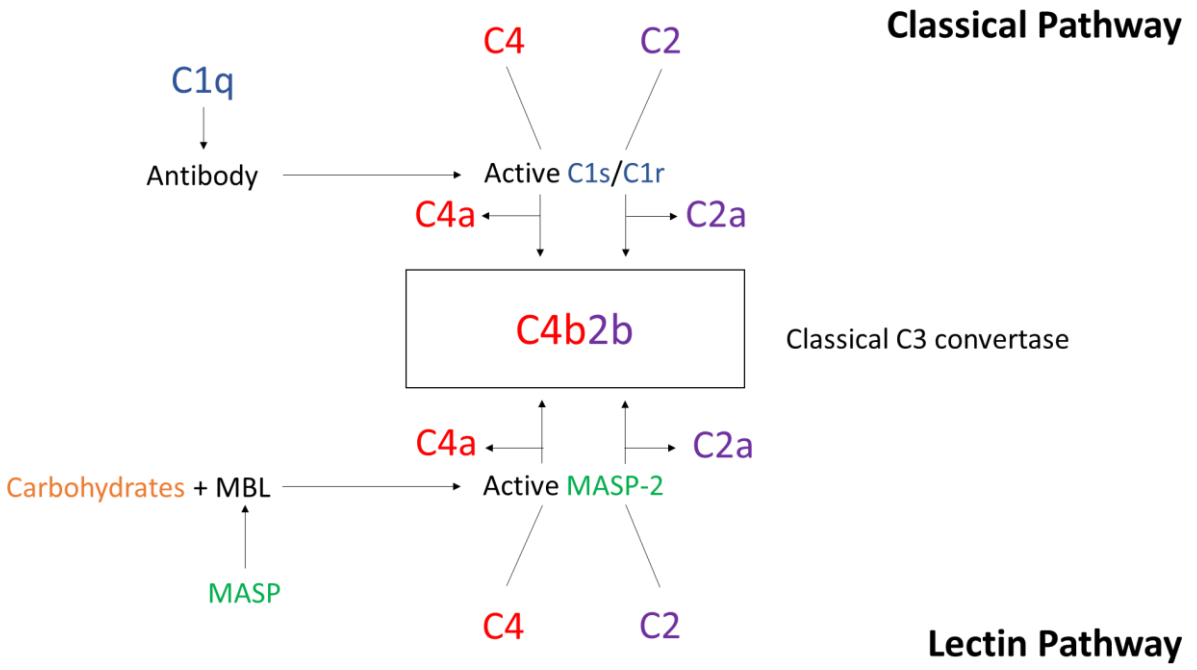


Fig. 1.6: a schematic of the classical and lectin complement pathways, leading to the formation of the classical C3 convertase.

The alternative pathway leads to the formation of the "alternative C3 convertase", C3bBb (fig. 1.7); this pathway is triggered by hydrolysed C3, $C3_{H20}$[140]. The hydrolysis creates new binding sites in C3, which is usually not active. The binding of factor B, a protease, to $C3_{H2O}$ leads to its cleavage by factor D to form a solvent-based alternative C3 convertase ($C3_{H2O}Bb$). This solvent-based alternative C3 convertase then cleaves C3 into C3a and C3b, and C3b then binds to the target cell membrane. The bound C3b then forms the alternative C3 convertase "C3bBb" by the cleavage of factor B by factor D.

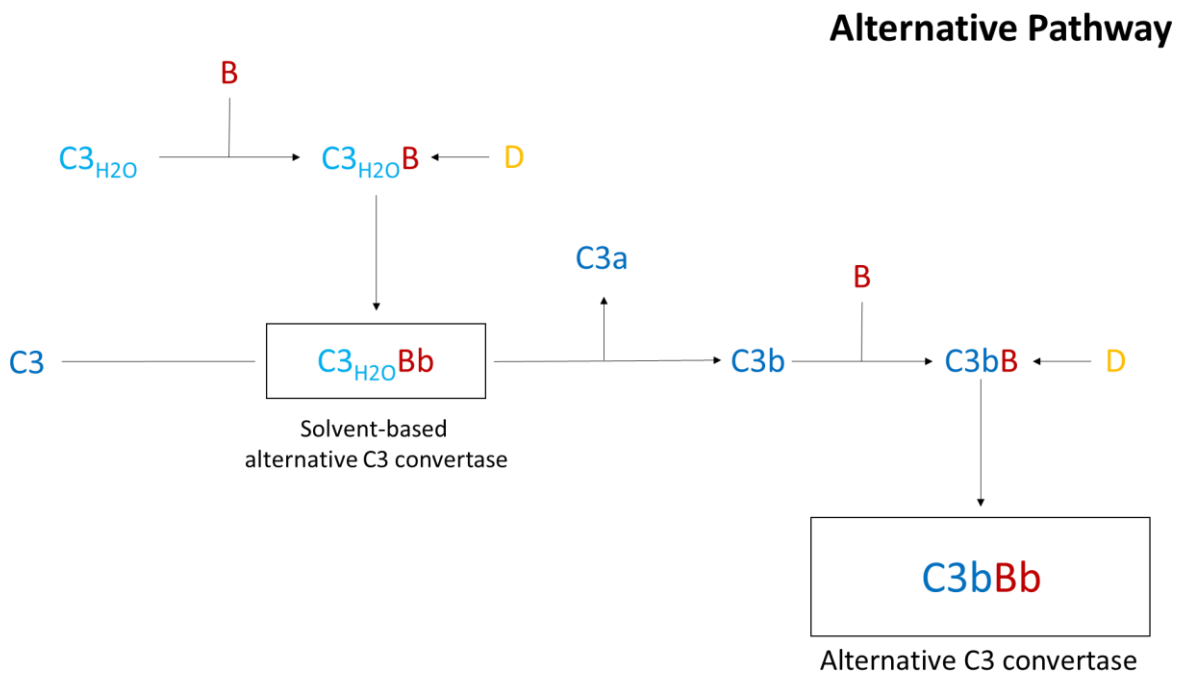


Fig. 1.7: a schematic of the alternative complement pathway, leading to the formation of the alternative C3 convertase by the action of a solvent-based alternative C3 convertase.

To counteract the effect of complement, *S. aureus* produces *Staphylococcus* complement inhibitor (SCIN) which inhibits the activity of the C3 convertase by stabilising the convertase[142]. The C3 convertases can dissociate which leaves C4B and C3b bound on the surface, and these act as co-factors for further C3 convertase formation[141]. This then amplifies the complement response, but SCIN inhibits this process and thus prevents the activation of complement[142]. Extracellular fibrinogen binding protein (Efb) was shown to interact with C3, therefore preventing C3 from attaching to the cell surface[143]. Staphylokinase binds to and activates plasmin, which then degrades IgG and C3b bound on the cell surface[144].

As mentioned above, C3a and C5a recruit immune cells by activating specific transmembrane receptors, and formyl peptides produced by bacteria also acts in the same way[140]. *S. aureus* produces chemotaxis inhibitory protein of staphylococci (CHIPS), which prevents chemotaxis by binding to C5a receptor and formyl peptide receptor[145]. Another protein produced by *S. aureus*, extracellular adherence protein (Eap)/MHC II analogue protein (Map), binds to ICAM-1 on endothelial cells, which then prevents LFA-1 from binding to ICAM-1[146]. This then means

that leukocytes cannot bind to endothelial cells, therefore cannot reach the site of infection by extravasation[140].

- *Preventing phagocytosis: Protein A, ClfA and capsule formation*

Protein A is a 45kDa protein found on the cell surface which is also secreted, and is expressed in most clinical *S. aureus* strains[147,148]. This protein contains 4-5 immunoglobulin binding domains, which bind to the Fc region of IgG and thus prevents the antibody binding to, or opsonising, its target antigen[140,148]. Antibody binding not only activates the classical complement pathway, but marks the antigen for phagocytosis, therefore a reduction in antibody opsonisation would lead to reduced phagocytosis[140]. Protein A also binds to the Fab region of VH3 idiotype immunoglobulins through a superantigen domain which interacts with CDR2[148,149]; this is capable of inducing B cell affinity maturation, however a study found that the response was limited and biased towards a VH3 antibodies. Therefore Protein A prevents the development of an antibody response to other *S. aureus* antigens[148].

The binding of fibrinogen by ClfA, which is expressed throughout *S. aureus* growth, can lead to the bacteria gaining a fibrinogen coat, and this may in turn protect the bacteria by preventing access to ligands which opsonising proteins would otherwise bind to; this has been shown to lead to reduced phagocytosis by murine macrophages[140,150]. Other fibrinogen binding proteins may also play this role, such as ClfB, which is more prevalent than ClfA during exponential growth[140].

Another mechanism used by *S. aureus* to prevent phagocytosis is the production of the polysaccharide capsule; most strains either express serotype 5, serotype 8 or serotype 336 capsules[151]. The capsule covers the surface of *S. aureus* cells, and serotypes 5 and 8 have been associated with increased virulence[152–154]. The capsule is thought to prevent opsonisation, and while complement can bind to the bacteria under the capsule, this is thought to be inaccessible to the complement receptors on neutrophils[140]. Capsule is however susceptible to antibody binding, and it has been found that high levels of anti-capsule antibodies gives protection from *S. aureus* infection by enabling the bacteria to be opsonised and then phagocytosed[155].

- *Resisting antimicrobial peptides and survival in neutrophils*

S. aureus naturally modifies its cell wall; Dlt includes D-alanine in ribitol teichoic acid and lipoteichoic acid, and MprF adds L-lysine to phosphatidylglycerol on the outer part of the cytoplasmic membrane[156,157]. These modification changes the charge of the cell, and makes it less susceptible to cationic defensins, which are secreted in the phagosome formed in neutrophils, and phospholipase D2 and lactoferrin[140]. Staphylokinase, mentioned above, has an additional role of binding to defensins, providing protection from this antimicrobial peptide[158]. Aurelolysin is an extracellular metalloprotease, and this is thought to cleave and inactivate cathelicidin LL-37, a defensin peptide[159].

To survive in neutrophils, *S. aureus* has developed the ability to interfere with endosome fusion and release of antimicrobial substances[160]. These are mediated through the production of factors regulated by SarA[140,161]. Staphyloxanthin, the yellow/orange carotenoid pigment which gives *S. aureus* colonies its distinctive colour, has been shown to bind oxygen free radicals which are produced during the oxidative burst during phagocytosis[162,163]. Other factors such as two superoxide dismutases and three methionine sulphoxide reductases and Mn^{2+} uptake has also been shown to remove oxygen free radicals[164–166]. Superoxide dismutases are enzymes which catalyse oxygen free radicals, and Mn^{2+} acts as a non-enzymatic superoxide dismutase[140,164,165]. Methionine sulphoxide reductases reduces the sulphur in methionine, which can become oxidised by oxygen free radicals and damage proteins[167].

- *Biofilm formation*

Another way which *S. aureus* evades the immune system is by forming a biofilm; an aggregate of bacterial cells encased in an extracellular matrix composed of polysaccharides, proteins and DNA (fig. 1.8)[129]. The process starts by the adhesion of *S. aureus* cells to a surface, where the cells then multiply. As this happens, an extracellular matrix is produced which encases the cells. Before the biofilm matures, there is an “exodus” where some cells are released from the biofilm. Those that have not dispersed then form a mature biofilm with secondary structures. During the final stage, cells detach from the biofilm causing it to disperse.

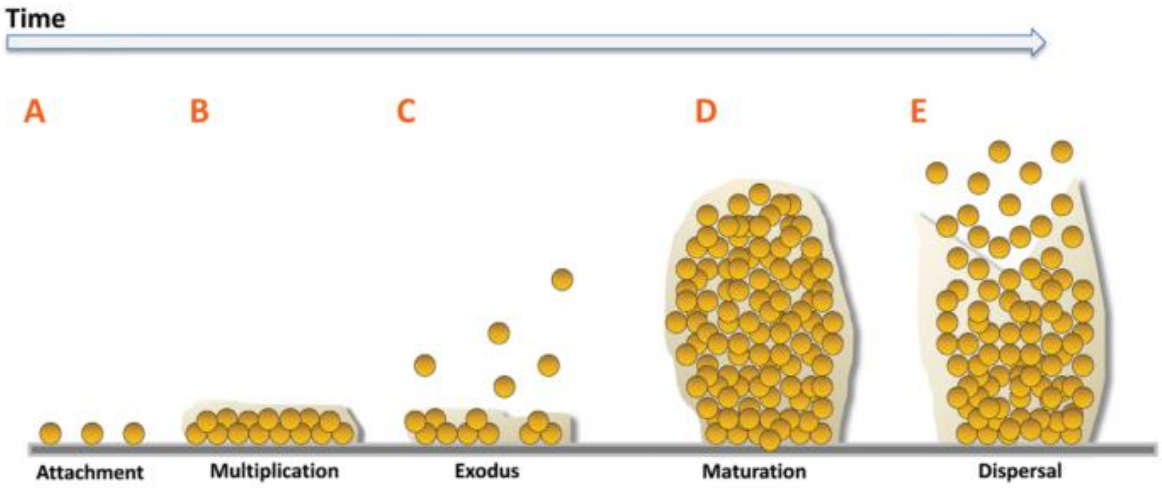


Fig. 1.8: The stages of biofilm formation in *S. aureus*, taken from Moormeier *et al*., Mol Microbiol 2017[129]. Biofilm formation begins when cells attach to a surface (A), where they then multiply (B). There is an exodus of cells (C) prior to the maturation of the biofilm (D). Finally, cells from the biofilm disperse (E).

During the attachment phase, where *S. aureus* attaches to host-derived extracellular matrix proteins, cell wall-anchored proteins play a key role; the MSCRAMMs enables the *S. aureus* to bind to host tissue but implanted materials which have been coated with host extracellular matrix, such as intravascular catheters. When it comes to abiotic surfaces, it is believed that adhesion is mediated by electrostatic and hydrophobic interactions and negatively charged teichoic acid rather than the adhesins mentioned above[129,168,169]. The autolysin AtlA also aids the attachment of cells to hydrophilic and hydrophobic polystyrene, and further evidence for this and cell wall-anchored protein independent attachment comes from a study using Nebraska transposon (Tn) mutants[170] with a Tn insertions in some of the cell wall-anchred proteins. These included *agrA* (part of the virulence regulating *agr* system), *atlA* and sortase A and B; only the *agrA* and *atlA* mutants had an effect on biofim formation *in vitro*, suggesting that cell wall-anchored proteins are not required[129].

Before the development of a mature biofilm, the adherent *S. aureus* cells multiply, and this occurs in the presence of sufficient nutrients[129]. In the absence of the extracellular matrix the newly formed cells can become detached from the attached cells, particularly if there are shear forces (i.e. in flowing fluid)[129]. Therefore, intercellular adhesion is important to ensure

that these new cells are not removed from the attached cells; cell wall-anchored proteins such as FnBPs, ClfB and SdrC are thought to have a dual role, in mediating intracellular attachment as well as attachment to biological matrix[171]. However, the role of other cell wall-anchored proteins in the accumulation stage are unclear, as some studies have suggested that they are involved while others have not seen a difference[172]. Other proteins seem to be involved in aiding the accumulation of cells during *S. aureus* biofilm formation, as protease treatment of multiplication-stage biofilms lead to abrogation of biofilm formation[172].

A study by Foulston *et al*. showed that enolase and GAPDH attached to the cell surface as the pH of the early biofilm decreased[173]. These proteins were not known to affect biofilm development, and lack an export signal meaning that there must be another mechanism by which they are released – the authors proposed that they are released by "regulated autolysis", in a similar mechanism by which extracellular DNA (eDNA) is released[173,174]. A study proposed that enolase and GAPDH bind to eDNA in low pH conditions, thus providing an extracellular matrix during early biofilm development[175]. Other studies suggest that phenol soluble modulins (PSMs), β-haemolysin, immunodominant antigen B (IsaB) also binds to eDNA and stabilises the early extracellular matrix[176–178]. Therefore, the combination of cytoplasmic proteins and eDNA may play a role in stabilising the early biofilm[129]. *S. aureus* is also capable of producing polysaccharide intercellular adhesin (PIA) through the *ica* genes, which forms a adhesive polysaccharide layer between the *S. aureus* cells, which then forms a part of the growing biofim matrix[179].

Prior to the maturation of the biofilm, there is a release of cells from the biofilm, and this forms the "exodus" phase of biofilm development[129]. This happens at the same time as the microcolony formation and causes the biofilm structure to be re-arranged[129]. This is distinct from the "dispersal" stage, and is triggered by the degradation of eDNA by nuclease rather than an *agr* response[129]. During this phase, only a small number of the biofilm population expresses the nuclease (*nuc*), which is regulated by the *S. aureus* exotoxin expression (*sae*) system[172]. Coagulase (*coa*) is another factor under *sae* control, and in the presence of host matrix proteins it is known to facilitate biofilm formation[180]. Notably, it is during this phase of biofilm formation that the biofilm integrity shifts from dependent on proteins during the

attachment phase to relying on both proteins and DNA[172]. The specific role of this "exodus" phase plays in biofilm maturation is unclear, however *nuc* mutants do not display microcolonies when they form biofilms, thus it is possible that the exodus of cells lead to re-structuring of the biofilm[172].

The formation of secondary structures, or microcolonies, increases the surface area of the biofilm and thus makes nutrient acquisition and waste removal more efficient[129]. Also, these microcolonies makes it easier for cells to disperse from the biofilm. This formation of microcolonies occurs during the maturation phase of biofilm formation. One possibility was that the biofilm was degraded by the release of PSMs, which are short peptide toxins and will be explored later on[181], forming channels, however time-lapse microscopy showed that these microcolonies form from cells which remain during the exodus, which forms the basal layer[129,182]. Studies observed two distinct types of microcolonies, a rapidly growing microcolonies expressing *lrgA* constitutively with delayed *cidABC* expression, and slow growing microcolonies constitutively expressing *cidABC* with no *lrgA* expression[183]. *lrgA* and *cidABC* operons are associated with cell-death, and the delayed *cidABC* expression seen in the fast growing microcolonies is thought to be induced by increasing hypoxia in these microcolonies[183]. A study revealed that *lrgAB* mutants had increased eDNA and adherence in the biofilm, suggesting that it functions as an inhibitor for *cid*-mediated cell lysis[184]. This would suggest that the slower growing microcolonies, which expresses *cidABC* without *lrgA*, would have a higher rate of cell death and thus increased eDNA.

Seemingly, during biofilm maturation *S. aureus* cells undergo metabolic diversification, and these two microcolonies displayed different dispersal rates, where "streaking" from a biofilm is seen in some but not all microcolonies[129,183]. Another advantage of having metabolically distinct populations could be advantageous for survival[129] – if one mode of metabolism leads to inability to survive in the environmental stresses, then there are cells displaying another mode of metabolism which may be better adapted for survival.

The final stage of biofilm formation is the dispersal of cells, and this is mediated by the *agr* system[185]. The *agr* system is a quorum sensing system which controls virulence in *S. aureus*,

and will be discussed further on[186]. *agr* deficient strains were shown to develop a stronger biofilm compared to wild-type strains[187]. Under low flow rate, AIP, the signal of the *agr* system, accumulates in the microcolonies and this seems to lead to biofilm dispersal[188,189]. One factor under control of the *agr* system are the PSMs. These toxins are expressed when AgrA, the response regulator of the *agr* system, binds to the *psm* operon promoters, and in the absence of these genes the biofilm formed was thicker compared to strains which harboured these genes[181,182]. PSMs are surfactant-like, and this is thought to disrupt the interaction of the biofilm matrix[129]. Alternatively, the PSMs could form insoluble aggregates promoted by eDNA, and this is thought to abrogate biofilm dispersal[178]. Therefore, the form that the PSMs take may play a role in biofilm integrity by maintenance of biofilm structure[129].

Toxins

S. aureus toxins can be grouped by their mechanism of action[190]. Cytotoxins such as α-toxin (also α-haemolysin) cause damage to cells by forming pores in the cell membrane, which in turn causes pro-inflammatory activity from the host. Other toxins include surfactant-like toxins and superantigens, which bind directly to the Major Histocompatibility Complex II (MHC II) proteins on T-cells, causing rapid proliferation of T-cells as well as a cytokine storm.

Toxins have a role in virulence by releasing nutrients and allowing the bacteria to invade further into the host or transmit to a new host[191]. Traditionally, it has been thought that high toxicity correlates with higher virulence, therefore leading to more severe disease – however, *S. aureus* switches off toxin production *in vitro*, as toxin production required a lot of energy. However, a study from our lab looking at a collection of USA300 MRSA isolates, from either carriage, SSTI or bacteraemia, and *S. aureus* isolates carried by a single patient, found that while high toxin production facilitates transmission of *S. aureus*, it was the strains which produced low levels of toxins which was correlated with disease severity (i.e. bacteraemia)[191]. These low-toxic isolates were shown to have better fitness in serum, thus have increased chances of establishing an infection in the bloodstream. Therefore, it seems that there is a trade-off between the ability to transmit between hosts (high toxicity) and maintaining fitness within a host (low toxicity).

It may be assumed that the action of these pore-forming haemolytic toxins on leukocytes may also aid in *S. aureus* colonisation, however these could also cause damage to the surrounding tissue and thus would induce an inflammatory response[192]. This would be counterproductive to long-term colonisation, however studies have found an inverse link between toxicity and degree of infection[191,193]. Therefore, it should be considered that *S. aureus* toxins may in fact be manipulating the immune system so that it can co-exist with the immune cells[192]. As Phenol Soluble Modulins (PSMs), which will be discussed in more detail later, do not act via a receptor, it could potentially target any cell, including other bacteria[194]. This could enable *S. aureus* to dominate its niche by inhibiting the growth of other bacteria, thus enabling *S. aureus* to access the nutrients found in the host.

- *Cytolytic toxins*

Haemolytic toxins are released as monomers, which oligomerise on the target cell membrane and cause the target cell to lyse[190]. Most of these toxins target immune cells, and the lysis of these cells are thought to have two affects. By killing the immune cells, *S. aureus* is able to evade phagocytosis as well as damaging the surrounding tissue by the pro-inflammatory response which occurs when the cytosolic components from the killed cells are released into the surrounding tissue[195]. Evidence has emerged to show that leukocidins may also interact with pro-inflammatory receptors, or have a "sub-lytic effect" as well[195].

α-haemolysin

The most well studied haemolytic toxin is α-haemolysin, or α-toxin, noted for its ability to lyse erythrocytes. It is released as a 33.2kDa monomer, which is water soluble and binds to target cells (erythrocytes. platelets, monocytes, lymphocytes, endothelial cells) and forms a heptameric barrel[196]. The ability of α-haemolysin to lyse erythrocytes means that *S. aureus* strains producing this toxin display β-haemolysis, where a clear zone of inhibition is visible on blood agar[197]. The process beings when the α-haemolysin monomer binds to the target cell membrane, which is able to interact with the lipid bilayer; at low concentrations, this is mediated by the ADAM10 receptor, a disintegrin and metalloproteinase[198]. The monomers diffuse in the membrane, which results in oligomerisation, then the formation of the heptameric barrel, which ultimately leads to the formation of a pore of around 1-2nm[190].

Formation of the pore leads to the lysis of the target cell through the efflux of small molecules and K^+ ions, and the influx of other small molecules, Na^+ and Ca^{2+} ions[190]. This disrupts the osmotic balance across the membrane, leading to rupture of the target cell.

The heptameric α-haemolysin is shaped like a mushroom of 100Å in height and up to 100Å in diameter, and the pore is a solvent-filled channel running through the middle (fig. 1.9)[196]. It was concluded from various studies that N-terminus of α-haemolysin has a role in opening the pore of the oligomeric toxin. The oligomeric structure of α-haemolysin can be subdivided into the cap, stem and rim domains; the cap domain, which along with parts of the rim domain protrudes from the target cell membrane, is composed of seven β-sandwiches. The stem domain forms the transmembrane component of the toxin, and a cleft is formed between the top of the stem domain and the rim domain; this cleft is rich in basic and aromatic amino acids, therefore could interact with the phospholipid head of the membrane.

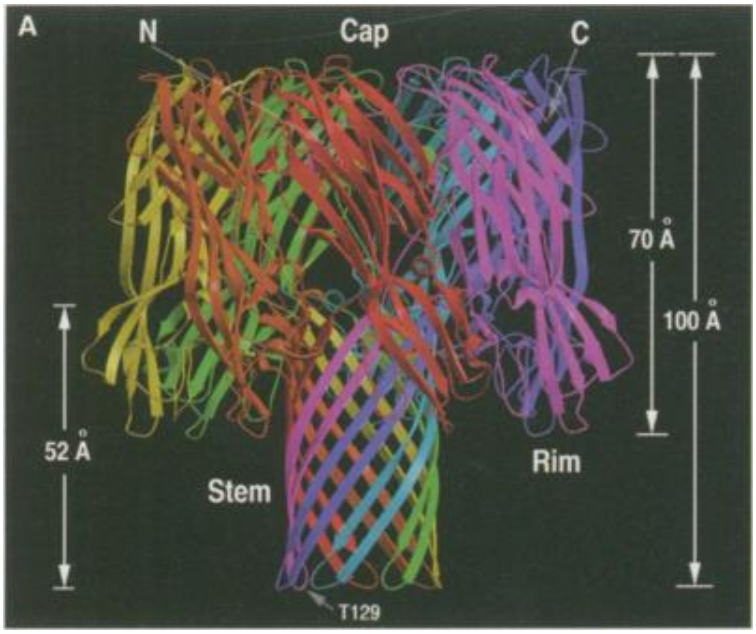


Fig. 1.9: the structure of heptameric α-toxin. Taken from Song *et al*., Science 1996[196].

Bi-component toxins

Some haemolysins are made of two distinct subunits, thus they are called "bi-component toxins", and this class of toxins includes γ-toxin (γ-haemolysin, HlgA, HlgB, HlgC), the leukocidins (LukED, LukGH (LukAB) and Panton-Valentine Leukocidin (PVL, LukSF-PV))[195].

These toxins are composed of an S subunit (HlgA/HlgC, LukE, LukH (LukB) and LukS-PV) and an F subunit (HlgB, LukD, LukG (LukA), and LukF-PV), and are structurally similar to α-haemolysin[195]. They are β-barrel pore forming toxins formed by the octamerization of four alternating S and F units (four of each). The subunits of these toxins are coded together on the genome with the exception of γ-toxin, which is organised as three open reading frames. These leukocidins also act via a receptor, including CCR5 (LukDE), C5aR and C5L2 (PVL), and CD11b (LukAB (LukGH)). The S subunit binds to the target receptor on the cell, which causes a conformational change thus allowing the F subunit to dimerise with the S subunit. The dimers then oligomerise to form the pre-pore, then the β-barrel transmembrane channel is inserted into the membrane to form the pore.

γ-haemolysin exists in two forms, HlgAB (classical γ-haemolysin) and HlgCB; HlgAB is mostly active against erythrocytes, while HlgCB is mainly active against leukocytes[199]. 89-100% of carriage strains have the genes encoding for γ-haemolysin, however studies have not been able to discover a link between γ-haemolysin and a particular infection[200,201]. However, the presence of HlgAB seems to promote survival of *S. aureus* in blood, most likely related to Fe^{2+} release from erythrocytes and macrophage evasion[199].

LukED has shown lytic activity against erythrocytes and leukocytes, however this is at a reduced level compared to γ-haemolysin[199]. As LukED also has affinity for a wide range of cells and host species, there has been animal studies carried out using this toxin; in a rabbit skin infection model, it was found that injection of LukED resulted in dermonecrosis, therefore it elicits a pro-inflammatory response *in vivo*[195]. It has also been shown that the presence of LukED is required for virulence in systemic infections in mice[202]. It is thought that LukED may play a role in impetigo, antibiotic associated diarrhoea, furuncles and invasive bacteraemia, as *S. aureus* isolated from these infections were found to contain the *lukED* genes more frequently[195]. However, since these studies it was found that *lukED* is present in a higher proportion of *S. aureus*, therefore the correlation between this toxin and the above infections may not be accurate[195].

LukGH, also called LukAB, was shown to be important in enabling *S. aureus* to survive in polymorphonuclear leukocytes; wild-type strains survived better than the *lukGH*/*lukAB* deletion mutant[195]. This was due to the damage caused to the polymorphonuclear leukocytes, as well as enabling *S. aureus* to escape from the phagosome; when bacteria are phagocytosed, they are restricted to the phagosome, however it seems that this toxin enables the *S. aureus* to escape from the phagosome and thus survive phagocytosis.

PVL is a phage-encoded bi-component leukocidin, and is frequently seen in CA-MRSA strains; these strains mainly affect people without risk factors for *S. aureus* infections and frequently causes SSTIs[4,199]. In contrast, the prevalence of PLV in carriage isolates is less than 3%[200]. It is thought that PVL+ve *S. aureus* strains have increased virulence, however this cannot be attributed to PVL alone[199]. These CA-MRSA clones are seen across the world, however it is thought that multiple *S. aureus* strains carrying a PVL phage and SCCmec VI (which causes methicillin resistance) expanded rather than a single PVL +ve MRSA strain[199]. In SSTIs caused by PVL+ve *S. aureus* strains, it has been demonstrated that PVL is expressed in the abscess, therefore PVL may be a contributing factor in the development of abscesses[199]. One study showed that PVL+ve *S. aureus* strains contributed to an increased ulceration and haemorrhage during pneumonia, likely caused by excessive inflammation triggered by the lysis of imunne cells by PVL[203]. However meta-analysis of epidemiological studies have not shown an association between PVL and increased virulence in *S. aureus* pneumonia[204].

β-haemolysin

β-haemolysin is a non-pore forming toxin, instead it is a sphingomyelinase C which hydrolyses the membrane lipid sphingomyelin into ceramide and phosphorylcholine and is thought to require Mg^{2+} ions as a co-factor[190,205]. While the precise mechanism of how β-haemolysin causes cell death is unknown, it could be due to the toxin causing changes in membrane fluidity of target cells, or by the formation of large ceramide-rich signalling platforms[205]. In the presence of DNA, β-haemolysin has been shown to covalently link with itself to form oligomers which then facilitates biofilm formation[177] - the initiation of a biofilm involves the attachment of *S. aureus* to a surface, and during this stage β-haemolysin produces an insoluble nucleoprotein matrix which becomes part of the biofilm matrix. This was also shown to

happen *in vivo*, where vegetation formed during endocarditis were found to be smaller when the infecting strain did not express β-haemolysin[177].

- *Phenol Soluble Modulins*

As the above toxins act through a receptor, they specifically target certain cells. *S. aureus* also produces toxins which do not act via a receptor, which are the surfactant-like PSMs[192]. These toxins are short, amphipathic peptides and are also produced by other staphylococci. They are grouped according to their size; the αPSMs (PSMα-1-4 and δ-toxin) are approximately 20-25 amino acids, βPMSs (PSMβ1 and 2) are approximately 44 amino acids[181]. They are encoded at different locations in the genome, and encode seven PSMs; the αPSMs genes are found in the *psmα* operon, as are the βPSMs (*psmβ* operon), and the δ-toxin is encoded within RNA III, the regulatory RNA effector of the *agr* response[181].

The charge of the PSM peptides seem to correlate with their cytolytic activity, with the αPSM displaying the highest cytolytic activity, δ-toxin displaying moderate levels and βPSMs are not cytolytic[181]. Studies such as the one by Periasamy *et al*. have shown that PSMs are involved in biofilm formation, particularly in forming secondary structures by forming the channels and microcolonies, and they are also important for biofilm dispersion as mentioned above[182]. αPSMs have been linked to increased virulence seen in CA-MRSA, while other PSMs have not been as clearly linked[206]. αPMSs have also been shown to prevent the inactivation of the antibiotic daptomycin in a study by Pader *et al*.[207]. Daptomycin binds to the most abundant phospholipid in *S. aureus* membrane, phosphatidylglycerol, and the study found that the release of phospholipids (mainly phosphatidylglocerol) bind to daptomycin which then renders the antibiotic inactive. However, it was only *S. aureus* strains lacking a functional *agr* system which inactivated daptomycin even though both this mutant and the wild-type strains used released phospholipids. The authors therefore looked for factors which are controlled by the *agr* system which could interfere with phospholipid-daptomycin binding and found that αPSM mutants were able to inactivate daptomycin. Therefore, αPSMs (particularly PSMα1) binds to phospholipids which means that they cannot bind to and inactivate daptomycin. Another study showed that PSMs are capable of regulating the expression of genes including the induction of genes necessary for the transport of PSMs[208].

- *Superantigens*

Superantigens, also known as pyrogenic toxin superantigens, include Toxic Shock Syndrome Toxin-1 (TSST-1) and staphylococcal enterotoxins SEA, SEB, SECn, SED, SEE and SEH[190]. Three functions seen in all these toxins are pyrogenicity (cause fever), and superantigenicity, and they have also been shown to increase the lethality of endotoxin in rabbit models[190]. These superantiens give rise to different toxin mediated infections, which will be explored below.

Toxic shock syndrome is an example of a superantigen mediated infection, and has been linked to the use of tampons during menstruation[190]. It is an acute systemic illness which can become fatal, characterised by high fever, rash, desquamation of the skin, low blood pressure and also involves other organs[209,210]. Early studies of patients with menstruation-associated toxic shock syndrome showed that they were not bacteraemic, therefore the syndrome was likely caused by factors produced by the bacteria[190]. TSST-1 was the first virulence factor to be associated with toxic shock syndrome, and is unique in its ability to cross mucosal barriers[190]. TSST-1 is the only toxin known to cause menstruation-associated toxic shock syndrome, while SEB and SEC have been seen to cause non-menstruation-associated toxic shock syndrome[190,211,212].

A property unique to enterotoxins, the other family of superantigens, is their ability to induce vomiting, hence it is a key factor in *S. aureus* food poisoning[190]. *S. aureus* food poisoning causes inflammation along the gastrointestinal tract, with the most severe lesions occurring in the stomach and the top of the small intestine[190,213]. When enterotoxin producing *S. aureus* strains colonise food, they can form toxins within the food, and ingestion of contaminated food results in vomiting and possibly diarrhoea, typically resolving in around 24-48 hours[190]. The incidences of *S. aureus* food poisoning is unclear, and as there are many enterotoxins involved prior infection may not result in protection from subsequent infections[190]. Although the mechanism of how enterotoxins induce vomiting has yet to be established, one proposed mechanism is that it is a response to inflammation caused by the enterotoxin[190].

Exoenzymes

S. aureus produces various enzymes, including nucleases, lipases and proteases[214]. These enzymes degrade host tissue, allowing the bacteria to invade deeper tissue, or degrade antimicrobial compounds produced by the host, such as antimicrobial peptides and fatty acids.

Aureolysin, a protease produced by *S. aureus*, inactivates PSMs, which then has an effect in the pathogenesis of osteomyelitis[215]. Aureolysin is also responsible for the activation of another *S. aureus* protease, SspA, and these along with staphopain A and B disrupts complement[216]. The pathogenesis of Staphylococcal Scalded Skin Syndrome is mediated by the cleavage of desmosomal cadherins by *S. aureus* exfoliative toxin, which is a serine protease[217]. Nuclease has a role in degrading neutrophil extracellular traps (NETs); these are DNA-based structures, with proteases antimicrobial peptides and histones bound to them, and these trap microbes[218]. The action of nuclease degrades these nets, thereby allowing the bacteria to escape from these NETs. Other enzymes, such as β-lactamases, actively degrade antibiotics to allow the bacteria to resist the action of the antibiotic, β -lactams in this case[98].

- *Lipases*

Lipases are enzymes which hydrolyse the ester bond in lipids, and *S. aureus* utilises lipases for lipid metabolism, and also as virulence factors[219]. Fatty acids are produced on human skin, and these can be harmful to *S. aureus*[220]. Lipases therefore break these down, so that these fatty acids are essentially disarmed. Lipases are also thought to be involved in the invasion of *S. aureus* into deeper tissue; this is supported by research which showed that *S. aureus* strains isolated from deep infections produce more lipase, and that lipase knock-out *S. aureus* lead to reduced bacterial load in a mouse model[221]. Also, lipases can be used as biotechnology catalysts, in industries such as food and pharmaceutical industries[222]. Chapter 5 looks at finding new genetic loci which regulate lipase activity in *S. aureus*, therefore more is discussed there.

Regulation of virulence factors

Virulence factors such as toxins are energetically costly to produce, therefore their expression is highly regulated[223,224]. These factors are generally either produced at low cell density (exponential phase) or high cell density (stationary phase); Protein A is an example of a protein expressed in the exponential phase, while toxins and enzymes are expressed during the stationary phase[223].

- *Accessory Gene Regulator*

One well studied regulator is the *accessory gene regulator* (*agr*) quorum sensing system (fig. 1.10)[186,223,225]. *S. aureus* strains with the *agr* system produces AgrD, which is processed by AgrB and a part of AgrD is exported through AgrB - this becomes the signal of this quorum sensing system, known as auto-inducing peptide (AIP)[225]. At high cell density, the bacterial cells recognise AIP through AgrC, a membrane-bound histidine kinase. AgrC then phosphorylates AgrA, the response regulator, and this causes transcription from the P2 and P3 *agr* promoters; P2 promoter transcript, RNA II, yields more Agr proteins (AgrD, AgrB, AgrC, AgrA) while the P3 transcript is a regulatory RNA, RNA III, the primary regulatory element of the *agr* system[225]. RNA III is thought to primarily act on transcription of the virulence factors, while acting on translation in some cases[225].

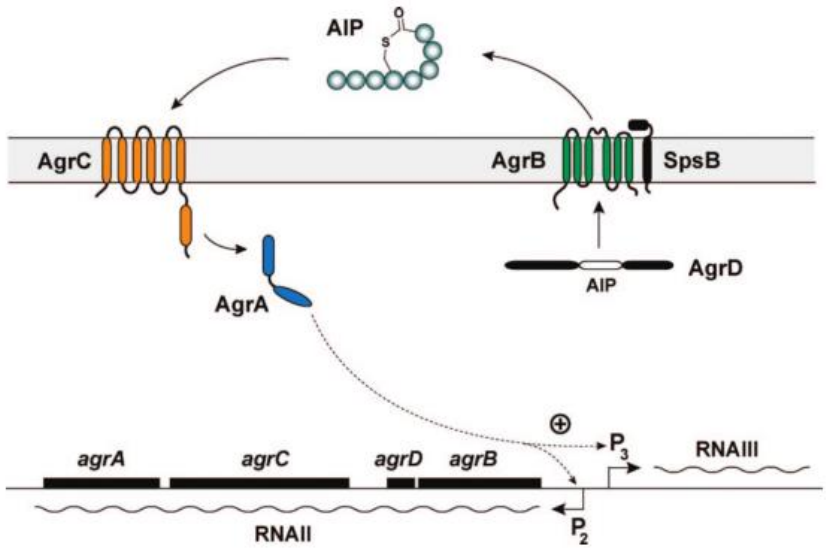


Fig. 1.10: a schematic demonstrating the auto-activating cascade of the *agr* system. Taken from Thoendel *et al*., Chem Rev 2011[225].

The *agr* locus may also be activated by the RNA III activating protein (RAP), by the RAP-TRAP two component system[223]. RAP is a secreted protein which is secreted throughout *S. aureus* growth, and it is proposed that it may phosphorylate TRAP that may lead to transcription from the P2 and P3 *agr* promoters. It is thought that AIP inhibits TRAP phosphorylation, thus it would be expected that RAP-TRAP and AIP do not activate the *agr* system at the same time.

The effect of *agr* on the genes it regulates vary depending on the kind of proteins which are encoded by the gene; for example, proteins promoting adherence to the host such as coagulase and fibronectin binding protein, are downregulated by *agr* activity while toxins and exoenzymes are upregulated by *agr* activity[223]. The upregulation of toxins by the *agr* system is thought to be the inhibition of repressor of toxins (Rot) by RNA III; Rot is a transcriptional regulator belonging to the SarA family, and downregulates the expression of toxins and exoenzymes while upregulating the expression of MSCRAMMs and Protein A. two loop-loop interactions between RNA III and *rot* mRNA prevents the translation of Rot, therefore changing the gene expression pattern.

The role of the *agr* system in pathogenicity has been established in animal models in studies such as the one by Abdelnour *et al*.[226]. However, clinical strains often have mutations in the *agr* locus, meaning that the activity is reduced, and there have been instances of clinical isolates which are *agr* negative[227]. However, for the USA300 lineage of CA-MRSA studies has found that *agr* activity is necessary for the development of SSTIs in a mouse model; one study found that mice infected with LAC, a USA300 strain, developed larger abscesses and showed necrosis in the epidermis and the dermis compared to an *agr* knock-out mutant or strain 252, a hospital-acquired MRSA strain[228]. The use of *agr* knock-out mutants and heat-killed bacteria in this study demonstrated that the development of larger abscesses with necrosis was due to toxins produced by LAC rather than an immune response, however comparison between LAC and 252 showed no difference in *agr* activity. Rather, it was found that there was a difference in the expression of PSMs, lipases and proteases seen between LAC and 252, which seemingly resulted in the larger necrotic abscesses.

Animal models have also shown that the *agr* system is required for the development of infective endocarditis, with mutants leading to vegetations with fewer bacteria compared to the wild-type[229]. It is interesting then, to note that *S. aureus* strains with an *agr* defect are frequently isolated from bacteraemia. Typically, mutations which inactivate the *agr* system occur mainly in *agrC* and *agrA*, and this has been shown to mainly occur prior to the onset of bacteraemia[230]. Therefore, it is plausible that selection for *agr* defective strains involve factors outside the bloodstream, such as those present in wounds or colonised intravascular

catheters. Interestingly, nasal colonisation does not seem to be a factor for the development of *agr* defective mutants, as majority of these *S. aureus* strains have a functional *agr* system[231]. However, this changes when the patients go into hospitals or are given antibiotics such as fluoroquinolones and β-lactams, therefore healthcare contacts seems to be a risk factor for the selection of *agr* defective *S. aureus* strains[231]. Studies have shown that *agr* system is not expressed in blood, even if there is a functional *agr* system in the strain, and this does not seem to be due to low density, as studies using higher densities have reported reduced RNA III expression in blood[232,233]. Apolipoprotein B is produced by the host in the serum, and this protein and other serum lipoproteins captures AIP as well as PSMs and thus disabling the activity of the *agr* system[232,234]. *S. aureus* can counteract this by aggregating, thus producing a higher local concentration of AIP, or by mutating *agrC* so that it is constitutively expressed[235,236]. Some cytotoxins such as α-haemolysin, γ-haemolysin and leukocidins are expressed in blood in an *agr* independent manner, and this seems to enable *S. aureus* to survive in the blood[230].

- *Staphylococcal Accessory Regulator A*

Another regulator of staphylococcal virulence is the *staphylococcal accessory regulator* A (*sarA*), which also controls virulence in a growth phase dependent manner[161,223]. This is regulated by σ^A -specific promoters, *sar*P1 and *sar*P2, as well as the *sar*P3 promoter, which has similarity to σ^B-dependent promoters. By being activated by different σ factors, the *sarA* operon is active during all growth phases; *sar*P1 and *sar*P2 during exponential growth and *sar*P3 during late exponential to stationary phase. As σ^B has a role in responding to stress, the *sar*P3 promoter is also activated during environmental stress.

The multiple regulatory mechanisms of *sarA* leads to peak expression during late exponential growth, although there is a possibility that the SarA protein may be under post-translational regulation as its levels remain relatively constant even when the *sarA* mRNA levels differ[223]. SarA is a DNA binding protein which binds to conserved AT-rich recognition sites, or Sar boxes, within the promoter of the target genes; one such binding occurs at the P2 and P3 *agr* promoters, thus altering the level of *agr* activity. Other genes regulated directly by SarA include Protein A, FnBPs, Cna, α-, δ- and γ-haemolysins. SarA binds as a homodimer, and its

binding reduces the number of base-pairs per turn of the DNA helix, which changes the spacing between the -35 and -10 regions within the promoter so that it either becomes optimal or becomes repressed. The pattern of regulation differs slightly between *sarA* and *agr*, as can be seen in table 1.1.

Gene	**+/- regulated by *agr*?**	**+/- regulated by *sarA*?**
cap8 (capsule)	+	+
spa (Protein A)	-	-
fnbA, *fnbB* (fibronectin binding proteins)	-	+
hla (α-haemolysin)	+	+
hld (δ-haemolysin)	+	+
hlgA, *hlgCB* (γ-haemolysin)	+	+
luk-PV (PVL)	+	+
lukED (leukocidin)	+	+
tst (TSST-1)	+	+
seb (enterotoxin B)	+	+
sspA (V8 serine protease)	+	-

Table 1.1: a table showing examples of genes which are either up-regulated (+) or down-regulated (-) by the *accessory gene regulator* (*agr*) or the *staphylococcal accessory regulator A* (*sarA*) regulatory systems. While the majority of these genes are regulated by the two systems in the same manner, there are some which are regulated differently. Based on tables from Bronner *et al*., FEMS Microbiol Rev 2004[223]

- *Staphylococcal Exoprotein Expression*

The staphylococcal exoprotein expression (*sae*) locus is another two-component regulatory system, like the *agr* system, and regulates expression of virulence genes mainly at a transcriptional level[223]. SaeS has high similarity with other sensor histidine kinases, and contains two transmembrane domains at the N-terminus and a histidine residue at the C-terminus which is auto-phosphorylated. SaeR is the response regulator which has similarities with other known regulatory proteins, and has a conserved aspartate residue which can be phosphorylated. The *sae* locus has been shown to be important in the expression of α- and β-haemolysin as well as Coa, but this is thought to be independent of the *agr* or *sarA* regulation. A study found that the level of *hla* (α-haemolysin) transcript was reduced in *sae* mutants both *in vivo* and *in vitro*, and that *hla* transcription was activated independently of *agr* or *sarA* in a

guinea pig infection model[237]. Another study found that the deletion of *saeS* lead to a loss of TSST-1 expression from a menstruation-associated TSS-causing strain, therefore the *sae* locus may play a role in the expression of this superantigen toxin[238].

Genome Wide Association Studies

Genome Wide Association Studies look for a link between genetic polymorphisms and a particular phenotype in a population - in bacteria this involves identifying SNPs and insertions and deletions (indels) within a population, then analysing to see if any of these genetic polymorphisms are statistically associated with a change in phenotype, such as toxicity[239]. The first GWAS were carried out for human diseases, such as Age-related Macular Degeneration, and the development of cheap whole genome sequencing technologies and the subsequent increase in the availability of bacterial genome sequences have facilitated the development of bacterial GWAS.

Success of bacterial GWAS requires a phenotype that can be tested, such as toxin production, a collection of sequenced strains, a statistical approach and then power calculations so that sufficient number of strains can be chosen[239]. Success of GWAS also depends on how the genetic polymorphisms are distributed in the population - the mixing of alleles within a population enables the identification of causal and passive loci. Causal loci are those which cause the effect, such as a change in toxin production, and passive loci are those that do not actually cause the effect despite being identified in the GWAS. Due to the fact that bacteria reproduce asexually, this leads to what is called a strong Linkage Disequilibrium (LD); LD is defined as "non-random association of alleles at two or more loci"[240], and in the absence of recombination or mutations a block of DNA will remain the same when passed onto the next generation. Genome variation does occur in bacteria, although not at each generation as is the case in humans; mechanisms include horizontal gene transfer, non-homologous recombination, homologous recombination and recurrent mutations. These distinct mechanisms lead to different branches of phylogeny containing similar genetic loci. Also, population structure is another important aspect which needs to be considered when designing GWAS. Due to these differences between human and bacterial genomes,

parameters in the tools developed for human GWAS needs to be adapted before carrying out bacterial GWAS.

Aims

The aim of this PhD was to see if functional genomics can be used to find novel ways in which virulence is regulated in *S. aureus*. Previous research in our lab has used GWAS to identify genetic loci which are associated with a change in toxicity in ST239 and USA300 MRSA strains, and this revealed that a SNP in the *ileS* gene was associated with a change in toxicity of these strains; this SNP is the most frequently seen SNP in isolates displaying low-level mupirocin resistance. The first part of the research focuses on how mupirocin resistance links to changes in toxicity. In the second part of the research, I applied GWAS to find out which genetic loci are associated with a change in lipase production of ST239 MRSA strains, as there is not much known about the regulation of lipase in *S. aureus*.

Chapter 2

- Materials and Methods

Laboratory Methods

Bacterial strains and growth conditions

Staphylococcus aureus strains were grown in Tryptic Soy Broth (TSB)/Agar (TSA) unless otherwise stated. *E. coli* strains were grown in Luria-Bertani (LB) medium. All were grown at 37°C, with shaking if using liquid medium.

To determine the mupirocin Minimum Inhibitory Concentration (MIC) of selected USA300 strains, 100µl of overnight cultures of these strains were plated onto agar plates containing the following concentrations of mupirocin; 0µg/ml, 32µg/ml, 64µg/ml, 96µg/ml and 125µg/ml and allowed to grow overnight at 37°C.

THP-1 cell growth conditions

THP-1 cells were grown in RPMI 1640 supplemented with foetal bovine serum (10%), L-glutamine (2mM), penicillin (100 units/ml) and streptomycin (100µg/ml). The flasks were incubated at 37°C with 5% CO_2, with sub-culturing every 2-3 days.

Polymerase Chain Reaction

DNA was extracted using High Pure PCR Template Preparation Kit (Roche); 5µl lysostaphin (5mg/ml)was added instead of lysozyme, and incubated for 30 mins. PCR was set up using GoTaq (Promega) OneTaq (NEB), DreamTaq (Thermo) or Phusin (Thermo) for cloning, and was set up according to the manufacturer's guidelines. The primers used are listed below (table 2.1).

The products were run on a 1% agarose gel in Tris-acetate-EDTA (TAE) buffer, at 85V for 35 mins, and visualised using a UV transluminator (BioRad).

- table 2.1a: PCR primers

name	**purpose**	**sequence**	**Tm (°C)**
ileS control	checks for the mupR SNP status	CTTATAAATTCTTACTTTCTCATGGTTTT	56.8
ileS S2	(mupirocin sensitive)	TAAATTCTTACTTTCTCATGGTTTTGG	57.4
ileS R4	(the mupR SNP)	TAAATTCTTACTTTCTCATGGTTTCT	55.3
ileS rev. 2		GATTGGTGCTAACAACTTCGTCATA	59.7
gyrB for	qPCR control primers (*gryB* gene)	CCAGGTAAATTAGCCGATTGC	57.9
gyrB rev		AAATCGCCTGCGTTCTAGAG	57.3
agrC RT fwd	*agrC* real-time PCR (qPCR)	GCAGATTATTCTATACTGTGCTAAC	58.1
agrC RT rev		ACTACAAAAAAGCTAGGGAATATTACAAA	58.2

- table 2.1B: cloning primers (for lipase project)

name	**purpose**	**sequence**	**Tm (°C)**	**restriction site**
SAUSA300_1966 cloing fwd	cloning of SAUSA300_1966	TTT GAATTC CTCCTTTTAAGATGTTTG	57.4	EcoRI
SAUSA300_1966 cloing rev		AAA GGTACC AAAGGAGCATAAAC	57.1	KpnI
murA cloning fwd	cloning of *murA*	TTT GAGCTC AAAACAAGATAAATCTATACA	58.6	SacI/BanII
murA cloning rev		AAA GGTACC AAGTAATGGACAAAGG	59.7	KpnI
atpH cloning fwd	cloning of *atpH*	TTT GAATTC TCACTCCTCTTTATAATTAATTAA	58.3	EcoRI
atpH cloning rev		AAA GGTACC GCGATAAATAATGGTA	58.1	KpnI

Creating and screening mupirocin resistant mutants

As *ileS* is an essential gene, it cannot be knocked-out. Therefore, we have created a selection based method where an overnight culture of the mupirocin sensitive strain was plated onto agar with 4µg/ml mupirocin. This was incubated for 42- 48h to allow colonies to grow. Few of these colonies were streaked out onto fresh agar with mupirocin (4µg/ml), and after these had grown fully (approx. 42h) a colony was inoculated into broth and grown overnight. This overnight culture was then mixed with glycerol and stocked in -80°C, or used to extract DNA (as above).

Once DNA was extracted, an allele specific PCR[241] was performed to determine what nucleotide is present at 1,762 of the *ileS* gene; G results in V588, which is mupirocin sensitive, and T results in F588, which confers resistance to mupirocin. The ileS control, ileS R4 and ileS S2 primers were paired with ileS reverse 2 primer, thus each colony was screened with three different primers; control to check for the presence of *ileS*, R4 to see if it is T at 1,762 and S2 to see if it is G at 1,762. The reactions were set up as recommended for GoTaq (Promega), and the cycling conditions were as recommended by the manufacturer; 50.5°C was used as the annealing temperature, and the extension time was 1 min. The products were visualised as above.

Genome sequencing using Illumina MiSeq was carried out by Prof. Sheppard and colleagues for the SH1000 mupirocin resistant mutants (MY40, SH1000:2, SH1000:3, SH1000:4, SH1000:5, SH1000:6), to check that this mutation in *ileS* had occurred, and also to ensure there were no other mutations in the genome which could account for a change in toxicity.

Toxicity assay

Strains were grown in a glass tube for 18h, and the supernatant was harvested by centrifuging up to 1ml of the above culture at maximum speed for 10 mins in a microcentrifuge. Meanwhile, the THP-1 cells were harvested by centrifuging at 1,200 rpm for 10 mins, then re-suspending in sterile phosphate buffered saline (PBS) to get a dilution of approx. $1*10^6$ cells/ml. 20µl of the bacterial supernatant was mixed with 20µl of THP-1 cells, and incubated for 12 mins at 37°C. 260µl of Guava ViaCount (Milipore) was then added to the sample, and

incubated at room temperature for 5 mins before analysing the viability count on the Guava flow cytometer (Milipore).

Growth curves

Mupirocin sensitive and mupirocin resistant strains were grown overnight, then diluted 1/500 in fresh broth before pipetting 200µl into a 90-well plate; the outer wells were filled with water. The lid of the plate was wiped with Triton X-100 before placing in a plate reader at 37°C, and OD_{600} readings were taken every 5 mins for 16h, with shaking in between each reading.

Analysing isoleucine content of mupirocin sensitive and mupirocin resistant cells

Cultures of SH1000 and MY40 were grown for 18h (5ml). The cultures were then washed 3 times with 1ml sterile PBS, centrifuging at 14,000rpm for 10 mins after each wash. The pellet was then re-suspended in 500µl sterile PBS, and 200µg/ml lysostaphin, 10µg/ml RNase A and 20µg/ml DNase I was added. This was then incubated at 37°C for 1h, and sonicated briefly on ice. The samples were then centrifuged at 14,000rpm for 10 mins, and 500µl of the resulting supernatant was applied to a Vivaspin 500 (3,000 MWCO) protein concentrator (GE Healthcare). This was then centrifuged at 15,000g for 30 mins.

The samples were analysed using Liquid Chromatography-tandem mass spectrometry (LC-MS/MS), similar to the protocol by Sowell *et al.*[242]. Five DL-isoleucine standards were made in 0.1% formic acid for quantification, and these were run before and after the samples through the system using a C18 column and positive ion mode MS/MS.

Making an agr knock-out strain

To eliminate the differential *agr* activity between strains, an *agr* knock-out of SH1000 was used (MY18). MY18 was created by phage transduction as follows; initially a φ11-ROJ48 lysate was created; ROJ48 was cultured overnight in LK (1% tryptone, 0.5% yeast extract, 1.6% KCl), and 200µl of ROJ48 overnight was mixed with 3ml LK and 3ml phage buffer (10mM $MgSO_4$, 4mM $CaCl_2$, 50mM Tris-HCl (pH 7.8), 100mM NaCl and 0.1% gelatine powder in molecular/MiliQ water) with up to 500µl φ11. This was placed on a shelf in a 30°C shaking

incubator until the media became clear. This 1st round lysates were then filter sterilised, and a second round of lysis was carried out with this first round lysate.

After these two lysis steps, the φ11-ROJ48 2nd round lysate was used to attempt a transduction into SH1000. Briefly, 1.8ml LK, 10µl 1M $CaCl_2$, 200µl of SH1000 overnight culture in LK and various volumes of the φ11-ROJ48 2nd round lysate (up to 500µl) was mixed together, then incubated at 37°C with shaking for 45 mins. 1ml ice cold 20mM trisodium citrate was added to the transduction mix, and incubated on ice for 5 mins. This was centrifuged to pellet the bacteria, which were then re-suspended with 1ml ice cold 20mM trisodium citrate. This was incubated on ice for 2.5h, and plated onto TSA plates with 20mM trisodium citrate, erythromycin and lincomycin (25µg/ml) which was incubated overnight at 37°C. One of the resulting colonies, SH1000 Δagr::P3-lux, was named MY18.

Transformation of pAgrC(his)A, a plasmid containing his-tagged AgcC and AgrA

pAgrC(his)A, a plasmid containing his-tagged AgrC and AgrA (AgrC(his)A), allows a strain harbouring it to respond to extracellular AIP without being able to produce AIP. This plasmid was extracted from *E. coli* TOP10 grown in broth with 100µg/ml ampicillin. This plasmid was extracted using a plasmid extraction kit (Thermo) following the manufacturer's protocol, and then electroporated into RN4220; briefly, RN4220 was grown overnight in Brain Heart Infusion (BHI) medium, and diluted 1:500 into 10ml fresh BHI. This was then incubated at 37°C with shaking until OD_{600} 0.4 - 0.6. The culture was then cooled, before centrifuging at 5,000rpm for 10 mins at 4°C. The media was discarded, and the pellet was re-suspended with 5ml ice cold 500mM sucrose. This was then re-centrifuged, supernatant discarded then 5ml ice cold 500mM sucrose added to the pellet. This above process was repeated for a total of 3 sucrose washes, and centrifuged as above.

After discarding the supernatant for the third time, 500µl ice cold 500mM sucrose was added to the pellet and re-suspended. This was left on ice for 30 mins, before centrifuging and re-suspending in 100µl ice cold 500mM sucrose. The cells were pipetted into an electroporation cuvette, and 5-10µl of the plasmid pAgrC(his)A was added on top. This was then electroporated using the StA setting on the MicroPulser (Bio-Rad), and then recovered in

750µl BHI at 37°C for 1h. 200µl of the transformant was plated onto agar plates with chloramphenicol, and incubated overnight at 37°C.

A mupirocin resistant variant of MY18, MY41, was created using the method described previously. pAgrC(his)A from RN4220 was electroporated into MY18 and MY41 as above, yielding MY42 and MY43. The resulting colonies were re-streaked onto fresh TSB-chloramphenicol plates, with mupirocin for colonies from MY41.

Protein extraction

-α-toxin and secreted Protein A 15ml of an 18h culture of SH1000 and MY40 were pelleted by centrifugation at 5,000rpm for 10 mins. The supernatant proteins were concentrated by precipitation using trichloroacetic acid (TCA). Briefly, 10ml of the supernatant was mixed with 2.5ml TCA and incubated on ice for 1h. This was centrifuged at 14,000 rpm for 20 mi.s at 4°C, and the pellet was washed 3 times with 300µl ice cold acetone. After the washes, the pellet was re-suspended in 80µl 8M urea and 0.5µl 1M NaOH.

-Cellular Protein A 15ml of an 18h culture of SH1000 and MY40 were pelleted by centrifugation as above and washed three times with phosphate buffered saline (PBS) (Difco). The pellet was re-suspended in 1ml PBS and lysostaphin (200µg/ml), DNase I (20µg/ml) and RNase A (10µg/ml) was added to the cells, and incubated at 37°C for 1h. The samples were then sonicated briefly on ice, and 700µl of lysate was mixed with 700µl PBS with 0.2% Triton X-100.

-AgrC(his) Overnight cultures of MY42 and MY43 was diluted 1:500 into 120ml broth with chloramphenicol, and grown for 2h at 37°C with shaking. 100nM AIP-1 was added to the culture to induce the expression of AgrC(his). This was allowed to grow for 1-3h, and 30ml samples were taken every hour from 1h after induction. Samples were centrifuged as before, media discarded and the pellet re-suspended in 100-250µl sterile PBS (100µl for 1h samples, 150µl for 2h samples, 200µl for 3h samples and 250µl for 4h samples). Lysostaphin, DNase I and RNase A was added to the cells as above, and incubated at 37°C for 1h. The samples were then sonicated briefly on ice, and mixed with equal volume of PBS with 0.2% Triton X-100.

-Phenol Soluble Modulins (PSMs) An overnight culture of SH1000 and MY40 was diluted 1:1000 into 50ml fresh media, and grown for 18h. The cultures were centrifuged at 18,000 rpm for 10 mins, and 35ml of the supernatant was mixed with 10ml butanol. The samples were incubated at 37°C shaking for 3h, and were then centrifuged at 3,000 rpm for 3 mins and 1ml of the upper organic layer was taken off. The samples were then freeze-dried overnight and then re-suspended in 160µl 8M urea.

SDS-PAGE and Western blotting

Concentration of the protein samples were measured using Bradford reagent, and the samples were made equal concentration. The samples were mixed with the same volume of Morris SDS-PAGE sample buffer, then incubated at 95°C for 5 mins (Protein A, PSM) or 60°C for 2 mins (membrane, AgrC(his)) before loading into 12% SDS-PAGE gels (unless otherwise stated). 5µl pre-stained protein ladder was added to a free well, and the gel was run at 120V for 1.5h, unless otherwise stated. Once run, the gels were washed in deionised water for 15 mins on an orbital shaker before staining with GelCode Blue Safe protein strain (Thermo) for 15 mins, again on the shaker. The gels were de-stained in deionised water overnight.

AgrC(his) Western blotting was carried out using HisProbe HPR Conjugate (Thermo). 15µl of the stained sample was loaded into the gels, and the gel was run at 200V for 1h. The proteins from the gel were transferred onto nitrocellulose membrane by wet transfer and blocked in 50ml 5% skimmed milk for 1h on a shaking platform or 4°C overnight. It was then washed 4 times with 15ml TBS-T (25mM Tris-HCl, 150mM NaCl (pH7.2), 0.05% Tween 20) on a shaking platform for 10 mins. The membrane was then incubated with 10ml of HisProbe working solution (2.5ml 10mM Tris-HCl and 150mM NaCl (pH7.4), 7.5ml TBS-T, 5µl HisProbe (4mg/ml), 0.25g BSA) for 1h on a shaking platform. The above wash process was then repeated twice, and the membrane was then placed onto 4ml SuperSignal West Pico Chemiluminescent Substrate (Thermo) working solution for 5 mins before visualising using an imaging system. Densitometry to determine the strength of the signal was carried out using ImageJ.

Protein A Western blotting was carried out using goat polyclonal anti-Protein A HRP conjugate (Abcam). 15µl of the stained sample was loaded into the gel. The proteins were transferred

and blocked as above. The membrane was washed 5 times in PBS, and incubated with the antibody-HRP conjugate (1:5000 in PBS) for 1h. The membrane was washed 5 times with PBS and developed and was visualised as above.

α-toxin Western blotting was carried out using anti-Staphylococcal a-toxin antibody (rabbit-whole anti-serum) and Protein G horseradish peroxidase (HRP) Conjugate (Invitrogen). The proteins were transferred as above and blocked overnight at 4°C. The membrane was washed 5x in PBS, the antibody was diluted 1:3000 in PBS and added to the membrane. This was incubated for 2h on a shaking platform. The membrane was washed 5x in PBS, the Protein G-HRP Conjugate was diluted 1:1000 and was added to the membrane and incubated for 1h on a shaking platform. This was visualised using Opti-4CN substrate kit according to the manufacturer's protocol (BioRad).

RNA extraction

As the protein samples from 1h after AIP-1 induction consistently showed differences, this time point was chosen to see if there is a difference in transcription of *agrC*. Overnight cultures of MY42 and MY43 was diluted 1:500 into 3ml fresh broth with chloramphenicol. This was grown for 2h, and induced with 100nM AIP-1 for 1h. 2ml of these cultures were mixed with 4ml RNA Protect Bacteria (Qiagen), and the RNeasy Mini Kit (Qiagen) was used to extract RNA following the manufacturer's protocol. Lysostaphin (200µg/ml) was added to Tris-EDTA buffer (Ambion), and this was added to the sample after the RNA Protect step before continuing with the protocol. when the RNA was extracted, Turbo DNA-free kit (Thermo) was used to remove genomic DNA from the RNA samples; 3µl Turbo DNase was added to the sample and incubated for 1.5h at 37°C, the a further 4µl Turbo DNase was added and incubated for 1.5h. 35µl DNase inactivation reagent was added to the samples to inactivate the DNase according to the protocol.

cDNA synthesis

RNA concentration of the samples were measured using Qubit RNA Broad Range kit (Thermo) and normalised before using QuantiTect Reverse Transcription Kit (Qiagen) to convert the RNA samples into cDNA according to the manufacturer's protocol. After adding the reverse

transcriptase, the samples were incubated at 42°C for 20 mins before raising the temperature to inactivate the reverse transcriptase.

Real-time PCR

Primers for *gyrB*, a housekeeping gene, was used alongside those for *agrC* to standardise transcript levels. ssoAdvanced SYBR Green Supermix (Bio-Rad) was used, using a standard curve of known genomic DNA concentrations for each primer set. 5µl of samples, standards and water were pipetted into the wells of a 96-well PCR plate. The supermix was added to water and primers according to the manufacturer's protocol, and 15µl of this mix was pipetted over the DNA samples. This was then placed into a qPCR machine, and run using the manufacturer's recommendation. The quantity of *agrC* cDNA was divided by the quantity of *gyrB* cDNA to get a ratio of *agrC* transcription levels.

Competition assay

Two pairs of strains, SH1000 and MY40, and MY18 and MY41, were co-cultured to detect fitness costs associated with mupirocin resistance. The strains were cultured individually overnight, then diluted 1,000,000x. 25µl of the mupS strain and 25µl of the mupR strain were pipetted into 5ml fresh broth, and grown at 37°C with shaking for 24h. The mixed culture was diluted 10,000,000x, and the lowest two dilutions were plated onto TSA and TSA with 4µg/ml mupirocin and incubated at 37°C. The resulting colonies were counted, and the number of colonies from the mupirocin plate was subtracted from the count from TSA. The Malthusian parameter was calculated using the following formula:

Ln (final density (colony forming units (CFU)/ml) / starting density (CFU/ml))

The Malthusian parameters of the mupS and mupR strains were compared using a t-test.

Lipase assay

Strains were grown and supernatants harvested in the same was as for toxicity assays. The substrates, *para*-nitrophenol butyrate (pNPB) and *para*-nitrophenol palmitate (pNPP) (Sigma), were made to 8mM concentration in isopropanol, and mixed with a buffer (50mM Tris-HCl

(pH 8.0), 1mg/ml gum Arabic, 0.005% Triton-X100) in a 1:9 ratio to create the assay mixes. A standard curve using *para*-nitrophenol (pNP) (Sigma) was created using the assay mixes, and 200µl of each dilution was pipetted into individual wells of a 96-well plate (Costar). 180µl of the assay mixes was pipetted into the remaining wells of a 96-well plate, and 20µl of the harvested bacterial supernatant was mixed into the wells in duplicates. The plate was placed in a FLUOstar Omega microplate reader (BMG Labtech) at 37°C, and a reading at 410nm was taken every 5 mins for 1h. The absorbance readings were converted to µM pNP released/min. using the standard curve.

Cloning

Once genes of interest were found by GWAS and verified by screening Tn knock-out mutants[170], they were cloned into the plasmid pRMC2[243] and transformed into the relevant Tn mutants. pRMC2 was extracted from *E. coli* DH5α using a plasmid extraction kit (Thermo). Genomic DNA from JE2, the strain which the Tn mutants are derived from, was extracted using the method stated previously. PCR was carried out using Phusion DNA polymerase with the HF buffer (Thermo) as recommended by the manufacturer. The cycle conditions used were also as recommended by the manufacturer; the annealing temperature was set to 5°C lower than the lower Tm for the primer pair, and the extension times were 30 sec.s for USA300_1966 and *atpH*, and 1 min. for *murA*.

The products were run on a 1% agarose gel as above, and if correct size were purified using a PCR purification kit (Thermo). The PRC product, which will be the insert, and pRMC2 were digested with the relevant restriction enzymes; 25.6µl of DNA was mixed with 3µl CutSmart buffer and 0.7µl of each restriction enzyme (all New England BioLabs), and was incubated in a 37°C water bath for 1.5h. The DNA was then run on a 1% agarose gel as above, and the correct bands were excised from the gel. The DNA was extracted from the gel slices using a gel extraction kit (Thermo). Alternatively, the digested products were purified using a PCR purification kit.

The double digested DNA was then quantified, and mixed in a 1:3 plasmid to insert ratio into a 20µl ligation reaction containing up to 100ng plasmid, 2µl Ligase Buffer and 1µl T4 DNA

Ligase (both NEB). This was incubated at 22°C for 1h, or the ligase volume was reduced to 0.5µl and incubated at 16°C overnight. Alternatively, molar ratios of 1:2, 1:3, 1:5 and 1:7 were calculated using NEBio Ligation Calculator (NEB), and a 10µl ligation reactions were set up, using 25ng plasmid, 1µl T4 ligase buffer and 0.5µl T4 ligase and incubated as above. As controls, double digested plasmid without the insert was also ligated. Further controls include 1A: single cut vector; ligated, 1B: single cut vector; un-ligated, 2A: single cut vector (other enzyme); ligated, 2B: single cut vector (other enzyme); un-ligated, 3A: double digest; ligated, 3B; double digest; in-ligated, 4; vector and insert; ligated.

The ligated plasmid and the no insert control were transformed into $CaCl_2$ competent DH5α cells; at this stage, uncut vector was used as a transformation control. To make these competent cells, an overnight culture was diluted 1/250 into 50ml broth and grown until OD_{600} 0.5 – 0.8. The culture was then placed on ice for 30 mins, and pelleted by centrifuging at full speed for 10 mins at 4°C. The pellet was washed with ice-cold sterile water, and pelleted again as above. The supernatant was discarded, and the pellet was re-suspended in 15ml of ice-cold 0.1M $CaCl_2$. This was incubated on ice for 1 - 2h, and pelleted again as above. The pellet was then re-suspended in 500µl of ice-cold 0.1M $CaCl_2$ with 15% glycerol, and 50µl were aliquoted into sterile Eppendorf tubes. These were then stored at -80°C.

The aliquot of $CaCl_2$ competent cells was thawed on ice, and was mixed with 5µl DNA (ligated plasmid, etc.) and incubated on ice for 30 mins. The cells were then heat shocked for 2 mins in a 42°C water bath. This was incubated on ice for 5 mins, and then 750µl of media was added. The cells were then incubated at 37°C to recover, and 100µl was plated out onto selective agar plates. The remaining cells were pelleted and most of the media was discarded. The cells were re-suspended in the remaining 100µl media, and plated out onto selective agar plates. These were then incubated at 37°C overnight.

If the transformation was inefficient, then electroporation was used with a similar protocol to above; using sterile dH_2O instead of 500mM sucrose. EcI setting on the MicroPulser was used here. The resulting colonies, if any, were streaked onto fresh agar plates. These were then grown in broth with ampicillin, and the plasmid was extracted as above. A small volume of the

plasmid was digested as above, as well as with only one restriction enzyme (1.4µl) – this gives a total of three reactions, with a double digest and two single digests. The digests were then run on an agarose gel as above, and if the correct insert was present the plasmid was electroporated into RN4220 as described above. The plasmid was then extracted from the resulting colony, remembering to add lysostaphin and incubating for 30 mins at 37°C, and checked by digesting as above (with the single digests). If correct, this plasmid was then electroporated into the correct Tn mutant. For these complemented mutants with the ligated plasmid, anhydrous tetracycline (50-400ng/ml) is required to try induce the expression of the inserted gene.

Perl Codes

- *Bootstrap.pl*

```
#!perl

use 5.14.2;
use warnings;
use strict;

#Assisted by Dr Sion Bayliss.

#Variables:
my$output = "Genes and Locus Tags ST239.txt"; #Output file. To be used in next script.
my$file = "S. aureus TW20 (ST239) full sequence.gb"; #The GenBank file.
my$line = ''; #Each element of the file.
my$locus_tag = ''; #Stores the locus tag.
my$genes = 0; #Count of the number of genes.
my$protein = ''; #Translated gene.
my$check = 0; #keep on reading file.

#Create the output file.
open OUTPUT, ">$output" or die "$output could not be created.\n";

#Open the GenBank file.
open FILE, $file or die "$file did not open.\n";

#Loop through the file, and print locus tags (with a >) and protein sequence into the output
file.
while (<FILE>){
	$line = $_;
		if ($line =~ /\/locus_tag\=\"(\S+)\"/){
			$locus_tag = $1; #$1 captures what's in the first bracket.
```

```
        } elsif ($line =~ /\/translation=\"(\S+)\"/){
                print OUTPUT ">$locus_tag\n";
                print OUTPUT "$1\n";
                $genes ++;
        } elsif ($line =~ /translation=\"(\S+)/){ #beginning of protein
                $protein = $1;
                $check = 1;
        } elsif ($check == 1 and $line =~ /(.+)"/){ #end of protein
                $protein .= $1;
                $check = 0;
                $protein =~ s/\s//g;
                print OUTPUT ">$locus_tag\n";
                print OUTPUT "$protein\n";
                $genes ++;
        } elsif ($check == 1 and $line=~/\s+(\S+)/){ #in-between
                $protein = "$protein$1"; #Another way of concatenating!
        }
}

#Print the number of CDSs found in the file.
say  "$genes CDSs found.\n";

#Close the files.
close OUTPUT;
close FILE;

exit;
```

- *Bootstrap 2.pl*

```
#!perl

use 5.14.2;
use warnings;
use strict;

#Variables for section 1:
my$file = 'Genes and Locus Tags ST239.txt'; #File from the previous program.
my$gene = ''; #the locus tag
my@sequence = (); #The protein sequence.
my$iso = 0; #Ile count for the gene.
my$isoleucine = 0; #Count of total Ile in the gene.
my$total = 0; #Count of total amino acids, including Ile.
my$gene_length = 0; #Lengths of each gene.
my@gene_list = (); #All the locus tags.
my@iso_summary = (); #All the Ile counts.
my@gene_lengths = (); #All gene lengths.

#Open the file with the locus tags and genes.
open INPUT, "$file" or die "$file could not be opened\n";

#Loop through the file, taking out the protein sequence to count the number of Ile.
while (<INPUT>){
	if (/^>(\S+)/){
		$gene = $1;
	} elsif (/(.+)/){
		@sequence = split (//, $1);
		$iso = 0; #Re-set count.
		$gene_length = 0; #Re-set count.
		foreach (@sequence) {
```

```
                        if ($_ =~ /I/i){
                                ++ $iso;
                                ++ $isoleucine;
                                ++ $total;
                                ++ $gene_length;
                        } elsif ($_ =~ /[ABCDEFGHKLMNPQRSTVWXYZ]/){
                                ++ $total;
                                ++ $gene_length;
                        } else {
                                die "Incorrect character $_";
                        }
                }
        push (@gene_list, $gene); #Names of genes (Locus Tags).
        push (@iso_summary, $iso); #Ile counts of the gene.
        push (@gene_lengths, $gene_length); #Gene lengths of the genome.
        }
}

#Print how many Iles there are and the total amino acids.
say "Out of $total amino acids, there are $isoleucine isoleucines.\n";

#Close the INPUT file.
close INPUT;

#Section 2 written with Dr Sion Bayliss
#Variables for section 2:
my$output2 = "Simulated Ile Distribution ST239.txt"; #Output - the distribution of the
simulated Ile counts.
my$output3 = "Random Ile 1M reps ST239.txt"; #Output - the data.
my@distribution = (); #Simulated Ile counts.
my$replicates = 1000000; #The number of times the simulation will be run.
```

```
my$simulated_isoleucine = 0; #Count of Ile in simulation.
my$number = 0; #Random number.
my$diff = ''; #Difference between observed and mean Ile.
my$low = ''; #Simulated Ile counts lower than actual.
my$high = ''; #Simulated Ile counts higher than actual.
my@max_min = (); #Sorted simulated Ile counts.
my$min = ''; #The lowest simulated count.
my$max = ''; #The highest simulated count.

#Create the two new output files.
open OUTPUT2, ">$output2" or die "$output2 could not be created.\n";
open OUTPUT3, ">$output3" or die "$output3 could not be created.\n";

#Print a header in OUTPUT3.
print OUTPUT3 "Gene Name\tGene Length\tObserved Ile\tMean Simulted Ile\tDiff in
Means\tMin. Simulated\tMax Simulated\tp-value\n";

#Simulating the Ile distribution.
for my$k(0..scalar (@gene_lengths)-1){ #Look at every gene in the genome.
        @distribution = (); #Clear the array.
                for (my$j = 0; $j < $replicates; $j ++){ #Repeat $replicates number of times;
                        $simulated_isoleucine = 0;
                for (my$i = 0; $i < $gene_lengths[$k]-1; $i ++){ #for each amino acid of the
        gene except the first,
#Generate random number between 1 and total amino acids-no. of genes (to get rid of first
Met).
                        $number = int(rand(($total-2779)))+1;
#If that number is below threshold (Ile in the genome), add to Ile count.
                        if ($number <= ($isoleucine)){
                                ++ $simulated_isoleucine;
                        }
```

```
                }
        push (@distribution, $simulated_isoleucine);
}

print OUTPUT2 "$gene_list[$k]\t";
print OUTPUT2 join "\t", @distribution, "\n";

my$m = 0; #Re-set the count.
foreach (@distribution){
        $m = $m + $_;
}

$m = $m / scalar(@distribution); #mean
$diff = $m - $iso_summary[$k]; #Difference in means.

#separate simulated Ile counts into those lower than actual and those higher than actual.
$low = scalar (grep {$distribution[$_] <= $iso_summary[$k]} 0..$#distribution);
$high = scalar (grep {$distribution[$_] >= $iso_summary[$k]} 0..$#distribution);

#calculating the P, = n (no. as extreme/more extreme than actual) / m (replicates).
my$p1 = '';
        if ($diff > 0){
                $p1 = $low / $replicates;
        } else {
                $p1 = $high / $replicates;
        }

#Sort the simulated Ile counts, to get the min. and max.
@max_min = sort {$a <=> $b} @distribution;
$min = $max_min[0];
$max = $max_min[-1];
```

```
#P where the actual is more extreme than all replicates.
if ($iso_summary[$k] < $min or $iso_summary[$k] > $max){
        $p1 = 1/$replicates;
        }

#Print the observed and simulated data into the OUTPUT3 file.
print OUTPUT3
"$gene_list[$k]\t$gene_lengths[$k]\t$iso_summary[$k]\t$m\t$diff\t$low\t$high\t$p1\n";
print "$gene_list[$k]\n";
}

#Close the output files.
close OUTPUT2;
close OUTPUT3;

say "\nAll Done!\n";

exit;
```

Chapter 3

- Verifying the Association Between Mupirocin Resistance and Toxicity in *Staphylococcus aureus*

Introduction

The increased risk of infections caused by carriage of MRSA has led to screening and decolonisation to reduce this risk in hospitalised patients, especially those about to have surgery. Most of these regimes include using the antibiotic mupirocin to decolonise the nose. Mupirocin is a polyketide antibiotic made by *Pseudomonas fluorescens* as a mixture consisting mainly of pseudomonic acid A (figure 3.1), and is used in ointment form to treat skin infections or to decolonise MRSA carriers[75]. Mupirocin acts by inhibiting the activity of isoleucyl-tRNA synthetase (IleRS), which charges tRNA with isoleucine (Ile), by mimicking isoleucyl-adenylate and binding to both the Ile site and the ATP-binding pocket; the 14-methyl terminus of the monic acid moiety imitates the side chain of Ile, and the pyran ring around C1-C3 binds to the ATP-binding site.

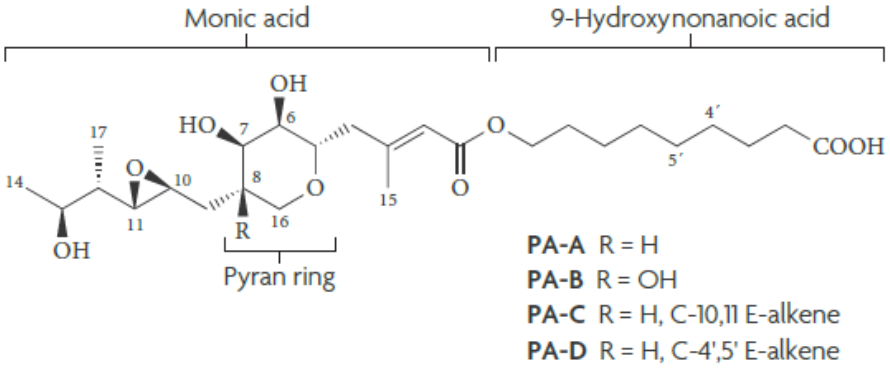


Figure 3.1: structure of mupirocin (pseudomonic acids), taken from Thomas *et al*., Nat Rev Microbiol 2010[75]. The variation in the R groups of the pyran ring gives rise to different pseudominic acids, i.e. pseudomonic acid A has H as the R group.

The inhibition of IleRS by mupirocin reduces the pool of isoleucyl-tRNA, signalling Ile starvation and thus triggering the stringent response. This is a conserved process which alters gene expression in bacteria to allow it to cope under stress[244]. In *S. aureus*, amino acid starvation leads to the production of ppGppp and pppGpp ((p)ppGpp), known as "alarmones" via the synthase activity of the bifunctional RSH (rel/SpoT homologue) enzyme[245]. This then leads to the re-programming of gene expression so that the available energy is focused into the necessary cellular processes, such as the synthesis of amino acid precursors. The end result of mupirocin inhibiting IleRS is that cell division does not occur, arising from the inability to synthesise new proteins due to a lack of charged isoleucyl-tRNA[244].

The use of mupirocin has led to resistance in *S. aureus*, which occurs at low-level (Minimum Inhibitory Concentration (MIC) 8-256µg/ml) and high-level (MIC ≥ 512µg/ml))[75]. Low-level resistance is conferred by Single Nucleotide Polymorphisms (SNPs) in the target gene, *ileS*, while high-level resistance is through the acquisition of an alternative IleRS, *mupA* or *mupB*[77]. Mutation of IleRS leads to changes in the structure around the Rossman fold of the protein, or the "synthetic site", where tRNA charging occurs, and this leads to mupirocin not being able to bind efficiently while allowing the enzyme to remain active. In *S. aureus*, there are three mutations in *ileS* which are commonly found; V588F, V631F and G593V[75].

The alternative IleRS is typically found on a plasmid, and is not inhibited by mupirocin, thus conferring high-level mupirocin resistance[75]. However, there have been reports of low-level mupirocin resistant strains carrying the *mupA* gene; in these cases, it seems that the *mupA* gene had become inserted into the chromosome, thus the strain is positive for *mupA* without having high-level mupirocin resistance[81,82]. This has implications in diagnosis, as these strains could potentially become highly resistant to mupirocin. Another report also found *S. aureus* strains carrying the *mupA* gene on the chromosome, however these strains retained the high-level mupirocin resistance mediated by *mupA*[83]. The other alternative IleRS is *mupB*, and is also found on a plasmid[77]. Despite having relatively low homology to *ileS* and *mupA*, the resulting protein contains conserved motifs found in class I tRNA synthetases.

Genome Wide Association Studies (GWAS) compare genetic variations against a phenotype to see which genetic polymorphisms are statistically associated with the phenotype[239]. One of the mutations conferring low-level resistance to mupirocin was found to be associated with toxicity in two distinct lineages of MRSA – USA300 and ST239 – through GWAS[246]. This particular association was shown to be an epistastic association, meaning that this single nucleotide polymorphism (SNP), a change from G to T at position 1,762 in the *ileS* gene (G1762T SNP in *ileS*), leading to the amino acid change V588F, interacts with multiple other SNPs at other loci in the genome to confer this effect rather than on its own (Fig. 3.2). The SNPs which were found to be in epistasis with the G1762T SNP in *ileS* in the USA300 and ST239 strains were different, so the only SNP in common was the one in the *ileS* gene. In this chapter, we sought to verify this association and to see if there is also an association between high-level mupirocin resistance.

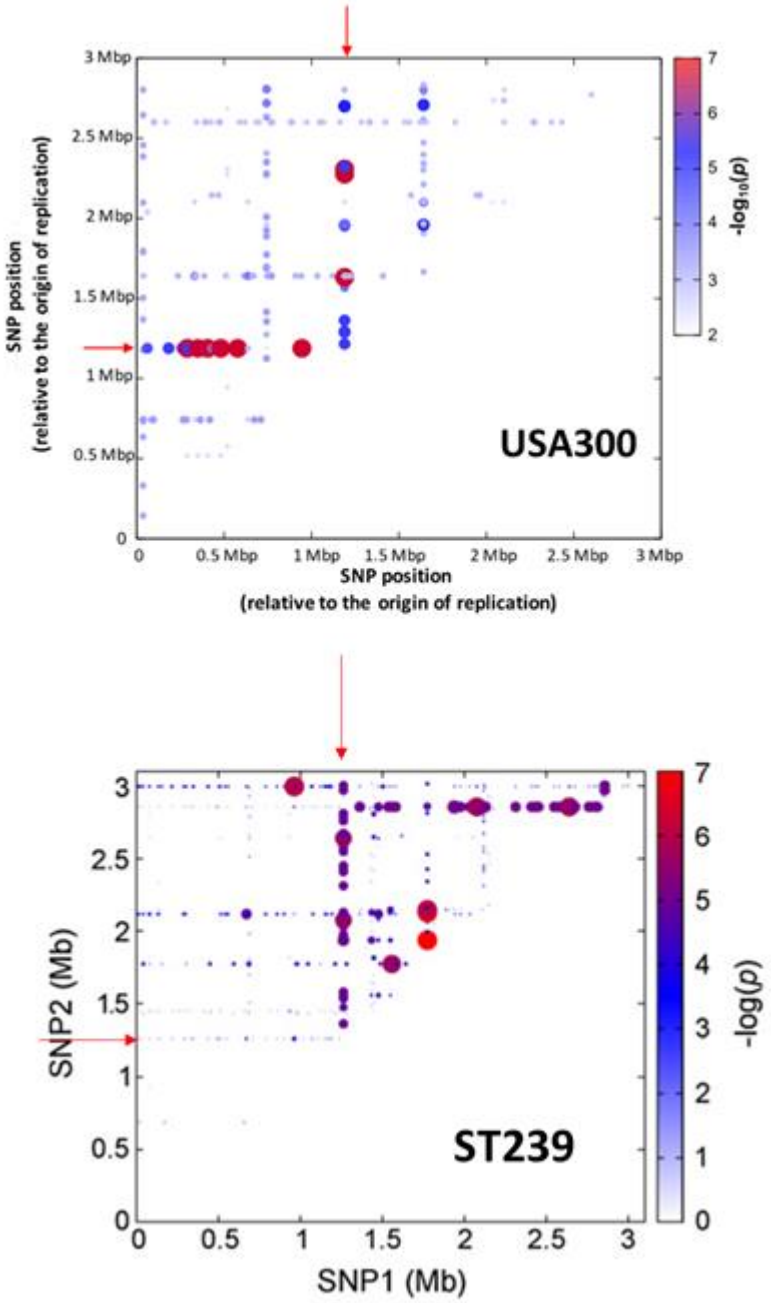


Fig. 3.2: heatmaps showing illustrating the epistatic interactions between the mupR SNP in *ileS* and other SNPs in USA300 and ST239 strain collections – the ST239 heatmap from Laabei *et al*, Genome Research 2014[246]. The axes represent the position in the genome, where 0 is the origin of replication. The circles represent instances where the presence of the two SNPs at these particular locations in the genome were associated with a change in toxicity. Larger, redder circles are where the interactions were more significant. The red arrows indicate the position of the G1762T SNP in *ileS* within the whole genome, and the circles represents the location of the second SNPs in the genome, and the presence of both are associated with a change in toxicity.

Results

Using genome sequencing data to find plasmids carried by the USA300 strains

Genome sequence readily reveals SNPs, as they reveal the positions where the nucleotide varies in different strains. However, detecting plasmids by sequencing is not as straight forward. In the USA300 lineage, pUSA03 is the plasmid which carries the *mupA* gene, so to find out which strains in our USA300 collection carry this plasmid collaborators at the Sanger Institute mapped the sequencing reads of our USA300 collection to known plasmids of the USA300 lineage (fig. 3.3)[247]. The data included the mapping to FPR3757 genome[248], which is the USA300 reference genome, so that we can have confidence that these strains are indeed USA300 and that the mapping is good, therefore the sequencing yielded good read coverage. Here, by using the mapping data provided by the Sanger collaborators, I was able to see if any of the sequencing reads matched known plasmids; for example, 8728_5#44 has a peak under "p18805-P03£, therefore I concluded that this strain carries the plasmid "p18805-P03". In this manner, I was able to separate our USA300 strains into those with and without each of these plasmids, with particular interest in pUSA03.

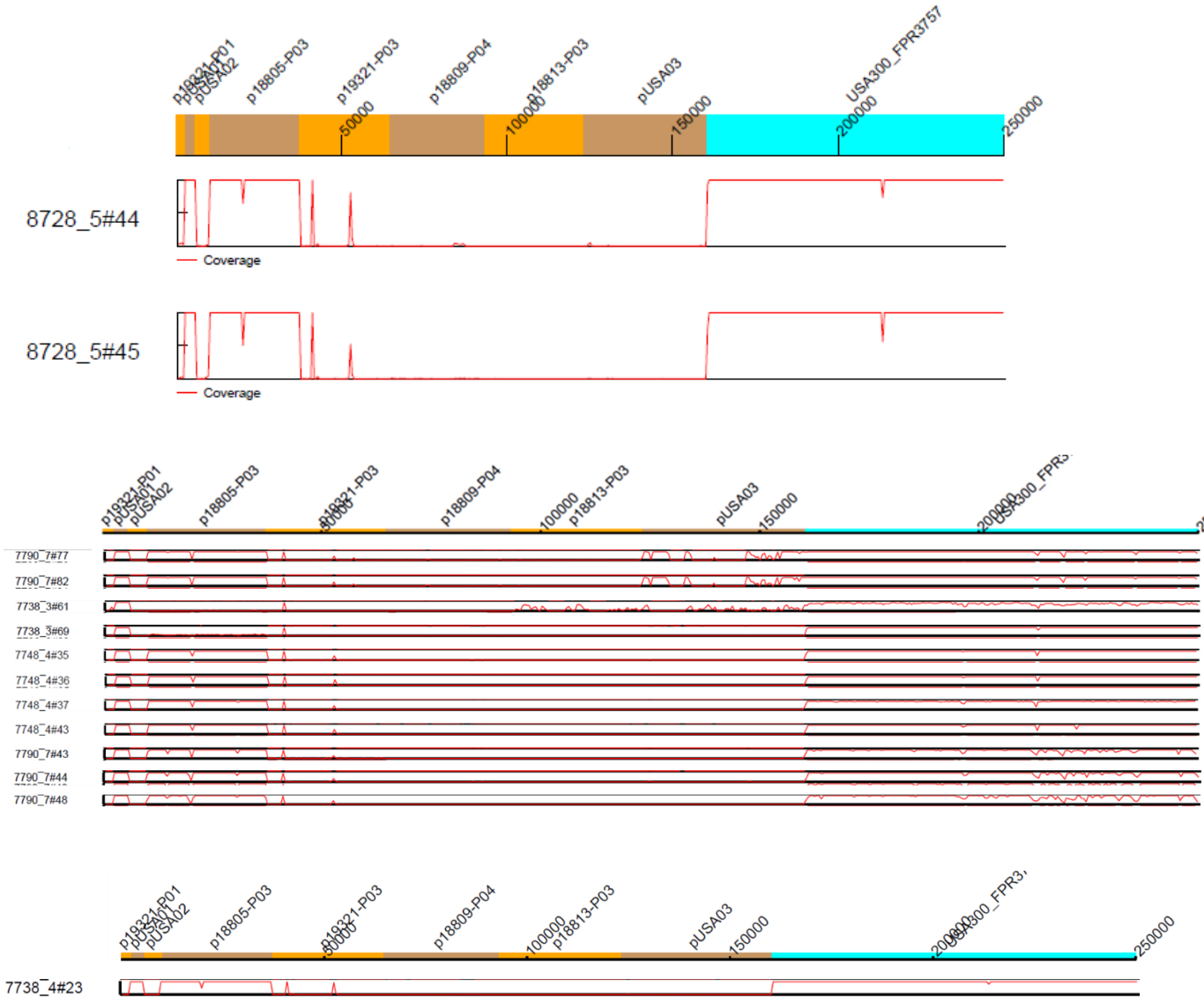


Fig. 3.3: genome sequencing reads from our USA300 collection were mapped to known USA300 plasmid sequences to provide a visual map, as shown in the above example, by collaborators at the Sanger Institute. This includes the USA300 genome on the right. Peaks represents areas where the sequencing reads matched to the sequence, therefore the presence of strong peaks under the plasmids shows that the strain carries this plasmid. By using this data, I was able to identify which USA300 strains carried plasmids.

G1762T Single Nucleotide Polymorphism in the ileS *gene and the* mupA *gene results in mupirocin resistance*

We next sought to verify that the G1762T in *ileS*, leading to V588F in IleRS, and the presence of *mupA* does result in resistance to mupirocin. The SNP will be referred to as "mupR SNP" from here on as none of the other low-level resistance conferring SNPs were present in this

collection of strains. To achieve this, we carried out minimum inhibitory concentration (MIC) assays on 14 USA300 strains; four are known to carry the mupR SNP and/or the plasmid carrying *mupA*, pUSA03, 10 being mupirocin sensitive. The mupR SNP was determined from the genome sequence of the strains, while the presence of *mupA* was determined by PCR, and confirmed by mapping sequencing reads to *mupA as described above*. The strains behaved as expected; the two strains with the mupR SNP only have low-level mupirocin resistance (mupR) (MIC between 64-96µg/ml), the strains with *mupA* have high-level mupR (MIC >128µg/ml) while the 10 strains without either are mupirocin sensitive (mupS) (table 3.1).

Strain	**MIC**	**mupR SNP?**	***mupA*?**	**PCR?**
7790_7#77	>128µg/ml	Yes	Yes	Yes
7790_7#82	>128µg/ml	Yes	Yes	Yes
7738_3#61	64-96 µg/ml	Yes	No	n/d
7738_3#69	32-64 µg/ml	Yes	No	n/d
7748_4#35	<1µg/ml	No	No	n/d
7748_4#36	<1µg/ml	No	No	n/d
7748_4#37	<1µg/ml	No	No	n/d
7748_4#43	<1µg/ml	No	No	n/d
7790_7#43	<1µg/ml	No	No	n/d
7790_7#44	<1µg/ml	No	No	n/d
8728_5#44	<1µg/ml	No	No	n/d
8728_5#45	<1µg/ml	No	No	n/d
7790_7#48	<1µg/ml	No	No	n/d
7738_4#23	<1µg/ml	No	No	n/d

Table 3.1: selected USA300 strains and their mupirocin Minimum Inhibitory Concentrations (MICs). The four mupirocin resistant isolates carried the mupR SNP, but two also carried the *mupA* gene. This is reflected in the MICs, where the strains with *mupA* have a higher MIC compared to those without this gene, and those without the mupR SNP were fully susceptible to mupirocin. This shows that the mupR SNP and *mupA* gene does confer low-level and high-level mupirocin resistance respectively

Does low-level mupirocin resistance affect toxicity of ST239/USA300 clinical isolates?

Having confirmed that the mupR SNP does confer low-level mupirocin resistance in the ST239 background we wanted to see if there were any differences in toxicity between mupS and mupR strains in our ST239 and USA300 clinical strains. Using toxicity data collected by a colleague (Dr Maisem Laabei) in the lab previously, we sought to see if there was a difference between mupS and low-level mupR strains of these MRSA isolates[191,246]. Dr Laabei used the

supernatant of ST239 strains to see how much lipid vesicles were lysed; these vesicles were designed to be sensitive to δ-toxin, Phenol Soluble Modulin (PSM)-α1, PSM-α2 and PMS-α3, and higher release fluorescence correlates with higher degradation of the vesicles. USA300 strains are a group of strains known to produce Panton-Valentine Leukocidin (PVL), so to test the toxicity of these strains Dr Laabei used THP-1 monocyte cells, which also have susceptibility to PVL. Then, I grouped the isolates into those with and without the mupR SNP and looking at the toxicity of those strains there was no difference in toxicity between mupS and mupR strains in either collection (p=0.4 for both USA300 and ST239 strains) (fig. 3.4). This is however expected, as the GWAS association was predicted to be due to epistasic interactions, therefore only apparent when present with specific combinations of SNPs in other genes.

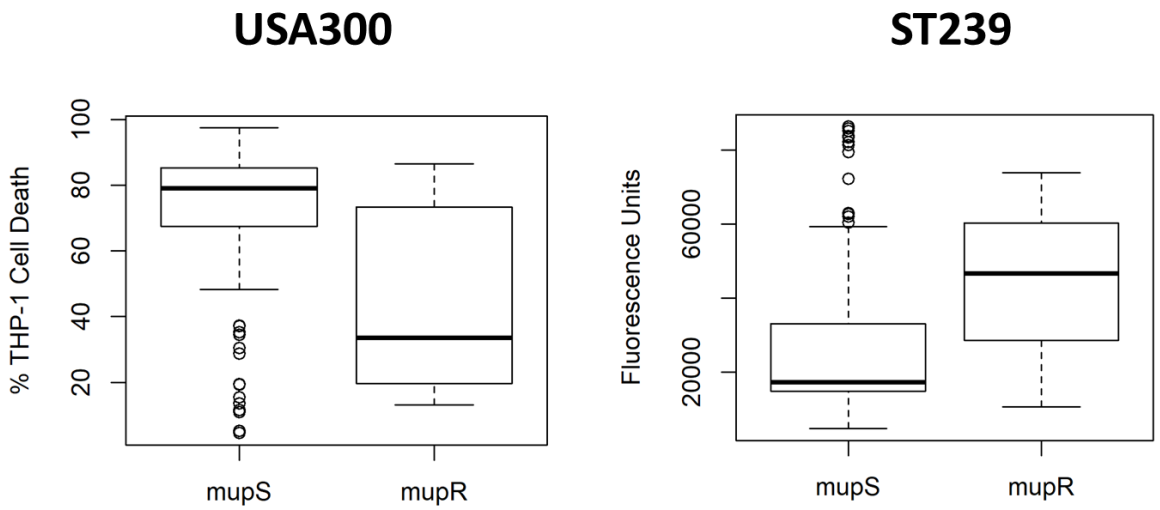


Fig. 3.4: Toxicity of USA300 and ST239 strains, using data collected by Dr Maisem Laabei. For both these groups of MRSA strains, mupirocin resistant strains show a difference in toxicity compared to mupirocin sensitive strains. However, these changes in toxicity are not significant (p=0.4 for USA300 both ST239, Kurskal-Wallis test), which was expected as this association was found to be an epistasis, therefore the presence of other SNPs at other loci are required. The box represents the lower and upper quartiles, with the line showing the median value. The whiskers show the lower and upper extremes, and the points outside the whiskers are the outliers.

High-level mupirocin resistance does not affect toxicity of USA300 isolates.

We next sought to investigate if this change in toxicity is also associated with high-level mupR; high-level mupR is conferred by the acquisition of the plasmid bourn *mupA* gene, also known as the alternative IleRS. Using mapped genome sequences from our Sanger Institute collaborators, we were able to identify which strains in our USA300 collection carried plasmids, as mentioned above. I then used this data to separate our USA300 strains into those with and without each plasmid, and the toxicity data collected by Dr Laabei was used to see if there was a difference between strains harbouring pUSA03 and those which do not. There was no difference in toxicity between strains with and without pUSA03 (p=0.4) (fig. 3.5, "pUSA03"). This was also true of other plasmids, so the carriage of the four known USA300 plasmids does not affect the toxicity of USA300 strains (fig. 3.5). As low-level mupR did reveal an association with toxicity in the GWAS studies, we chose to investigate this further.

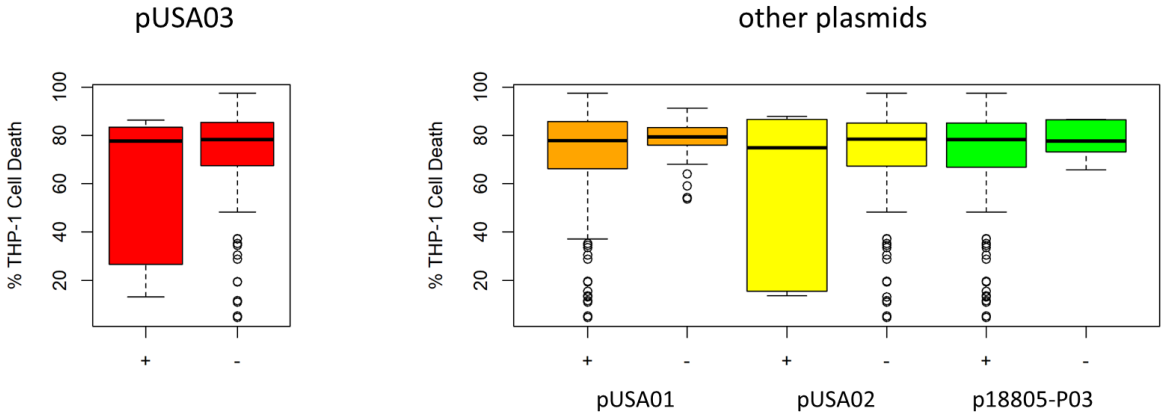


Fig. 3.5: Toxicity of USA300 strains with (+)/without (-) various plasmids. pUSA03 is the plasmid which contains *mupA*, the alternative IleRS, and there is no difference in toxicity between strains without and with this plasmid (p=0.4, Kruskal-Wallis test). Looking at the presence or absence of other known USA300 plasmids within the collection also show no differences in toxicity (p>0.05, Kruskal-Wallis (pUSA01)/t-test (others)). Therefore, the carriage of these four plasmids is not linked to a change in toxicity, even in the case of the *mupA* containing pUSA03. The box represents the lower and upper quartiles, with the line showing the median value. The whiskers show the lower and upper extremes, and the points outside the whiskers are the outliers.

Isolating and verifying mupirocin resistant SH1000 strains using a PCR based screen

Next, we sought to see if it was possible to select for this mutation in the *ileS* gene by growing strains in the presence of mupirocin. As *ileS* is an essential gene, it cannot be knocked out, therefore selecting for this mutation using mupirocin is the practical way of mutating this gene. Firstly, overnight cultures of mupS USA300 strains were grown without mupirocin, and then 100µl of this culture was plated onto agar plates with 4µg/ml mupirocin. After approximately 48h, we observed individual mupR colonies growing on some of these plates. To quickly analyse if these colonies had the desired mutation, we designed a PCR based screen; by designing primers in such a way that the T1762 would have one mismatch while the G1762 would have two mismatches, it was possible to differentiate the mupR SNP (fig. 3.6 A)[241]. The one mismatch does not affect the PCR, while two mismatches does, therefore no product is seen if the template DNA has G1762 while a product is seen if the template is T1762 (fig. 3.6 B, "Sens." (G) and "Res." (T)). This PCR based screen showed that this mutation had indeed occurred in these colonies (fig. 3.6).

To ensure that the change in toxicity is only related to the mupR SNP, we applied the same selection method on the lab *S. aureus* strain SH1000 and obtained mupR colonies; having isogenic mupS and mupR strains ensures that there are no other variations in the genome, particularly at the loci identified as being in epistasis with the mupR SNP by the GWAS. The colonies were analysed using the PCR based screen, and their genomes were sequenced by collaborators to ensure that there were no other mutations which could account for a change in toxicity. This shows that the method is effective for *S. aureus* strains from different lineages.

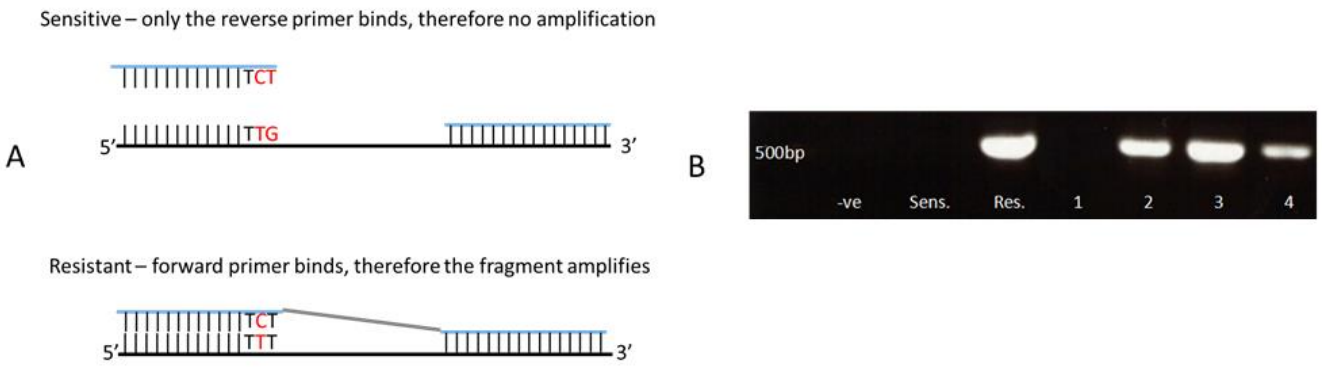


Fig. 3.6: Mechanism of the PCR-based screen for the mupR SNP; the design of the primers creates two mismatches if the strain is mupirocin sensitive, but only one mismatch if the strain has the mupR SNP (A). This means that if the strain does not contain the mupR SNP there will be no amplification (B - sens. and 1). However, there is only one mismatch if the strain has the mupR SNP, therefore there is amplification (B - Res. and 2-4). This result show that the parent strain (SH1000) is sensitive to mupirocin and does not contain the mupR SNP, while the SH1000-mupR mutants (2-4) contain the mupR SNP.

Does making SH1000 mupirocin resistant change its toxicity?

These mupirocin resistant colonies were then grown for 18h without mupirocin, alongside the mupS parent strain (SH1000), and the supernatant was analysed for toxicity; it was found that mupR had a small but statistically significant lowering effect on the toxicity of the resistant strain ($p=0.02$) (fig. 3.7).

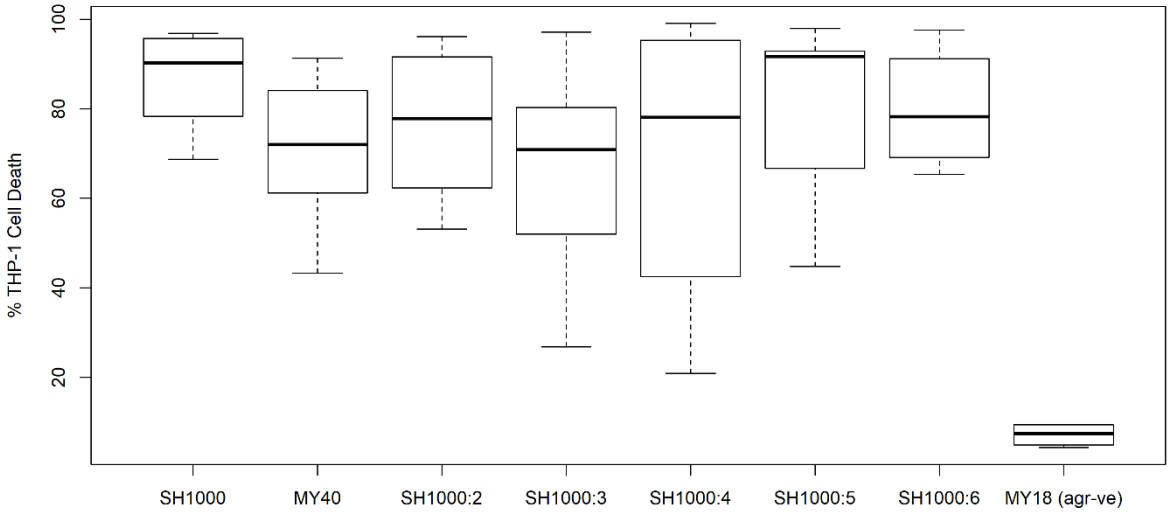
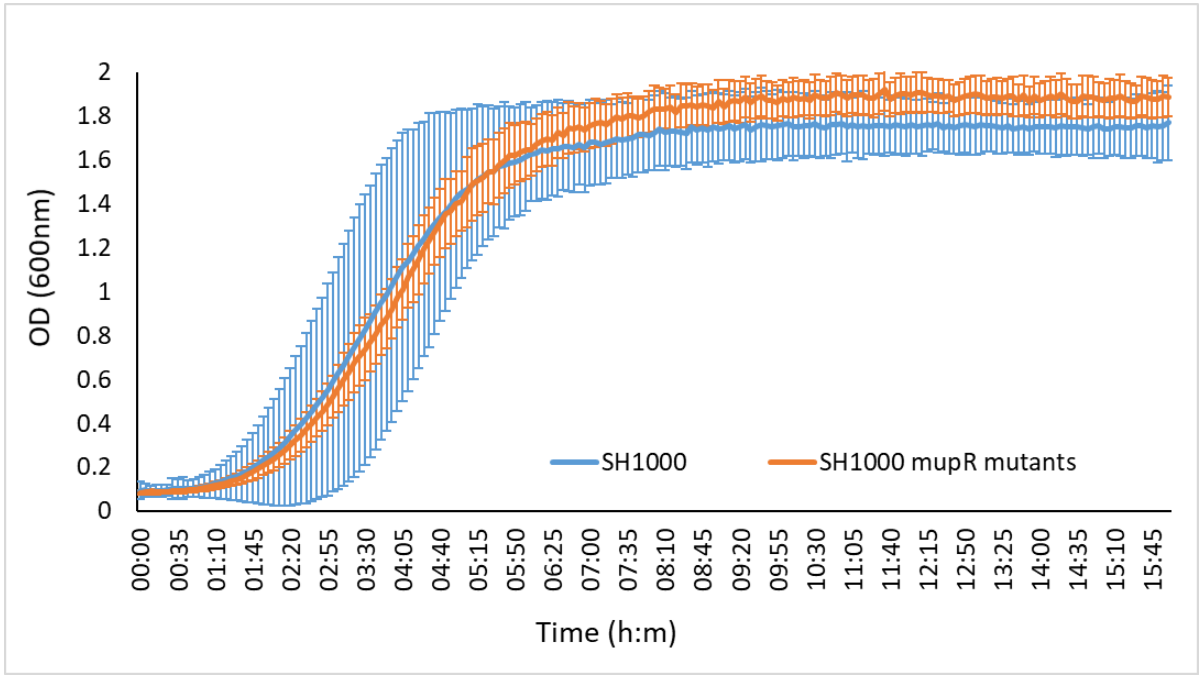


Fig. 3.7: toxicity of mupS, mupR and *agr*-ve SH1000 strains. There is a reduction in toxicity seen in most mupR strains (MY40, SH1000:2 - :6) compared to its mupS parent strain (SH1000) (p=0.02 for MY40, one-way ANOVA) (n=9 for mupS and mupR samples, and n=4 for MY18). However, this reduction is not as great as seen in the *agr*-ve strain (MY18) – the *agr* system is a global quorum sensing system which regulates many phenotypes including toxicity. The box represents the lower and upper quartiles, with the line showing the median value. The whiskers show the lower and upper extremes, and the points outside the whiskers are the outliers.

Does mupirocin resistance affect the growth of S. aureus*?*

Many antibiotic resistance conferring mutations have effects on the fitness of the bacteria, and noting that *ileS* gene is an essential gene which is highly conserved across *S. aureus*, it is surprising that the mupR mutation in this gene is reported to have no effect on the fitness[249]. In the presence of mupirocin, these mupR strains indeed grow slower than in the absence of mupirocin, taking around 48h to grow. However, comparing the growth of SH1000 (mupS) and the mupR mutants in the absence of mupirocin shows no difference in growth rate (fig. 3.8). This is in accordance with previous research, which found that this mutation is "cost free"[249]. Also, the fact that the growth rate of these strains do not differ would suggest that *agr* activity between them would most likely be similar, thus the difference in toxicity seen in these strains are not likely related to differences in growth.



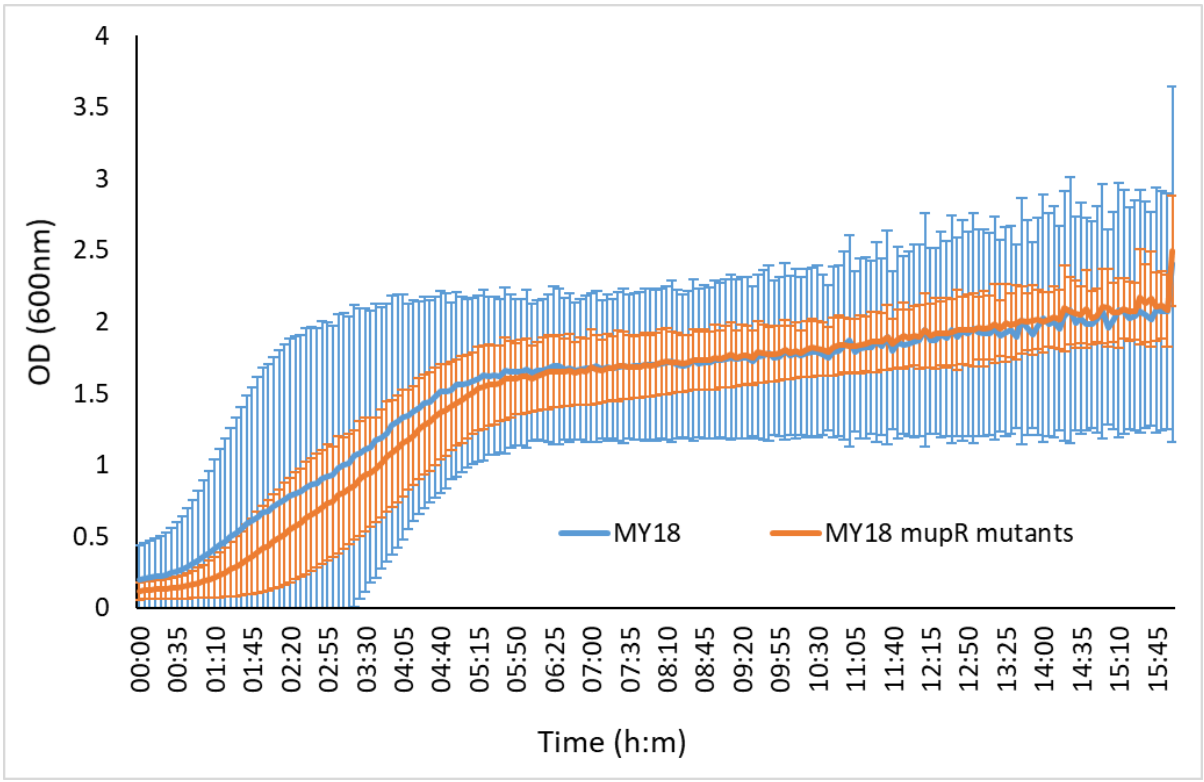


Fig. 3.8: growth of mupR strains compared to its mupS parent shows no difference (n=3 for SH1000 and MY18, n=6 for mutants). Both strains were grown as a pure single culture for 16h at 37°C, with OD_{600} readings taken every 5 mins. There is no difference between the growth of SH1000 (mupS) and its mupR mutants, therefore in pure culture the mupR SNP does not seem to affect fitness. The error bars show the 95% confidence interval.

Discussion

Mupirocin is a clinically important antibiotic, but growing resistance could hamper its use as an MRSA-decolonising agent. Mupirocin resistance (mupR) is seen in two levels, low-level and high-level, and a SNP in *ileS* mediating low-level mupR was found to be associated with a change in toxicity in two distinct lineages of MRSA – ST239 and USA300. Due to the high false-positive rate seen in bacterial GWASs, we sought to verify this association.

Initially, we sought to verify that the mupR SNP and the *mupA* gene does indeed lead to mupR; screening USA300 isolates with the mupR SNP, we found that they exhibited the appropriate resistance phenotype. With *ileS* being an essential gene, it is impossible to obtain viable knock-out mutants. However, as the SNP does lead to mupR, this could potentially be used as a selective pressure to obtain mupR mutants. Indeed, this process does work, and for the purposes of quickly detecting whether the correct mupR SNP was the cause of gaining mupR, we developed a PCR based assay. This approach could be applied to other antibiotic resistances resulting from SNPs, to develop a relatively fast screening method – this includes rifampicin resistance, where we did manage to get a similar assay working for one of the resistance SNPs.

The genomes of the USA300 and ST239 isolates had been sequenced, therefore those carrying the mupR SNP were quickly identified. Using toxicity data collected by a colleague (Dr Maisem Laabei), we were able to look for differences in toxicity between mupS and mupR isolates; indeed, we saw that toxicity was different in mupR isolates. From genome sequencing, it is possible for us to identify isolates carrying previously known plasmids; this was done by using the mapped sequencing data for the USA300 isolates provided by the collaborators at the Sanger Institute. This meant that we were able to separate isolates into those with and without plasmids, including pUSA03 which has *mupA*. When analysing the toxicity data, we saw no difference in toxicity between isolates carrying any plasmid, therefore the difference in toxicity observed in the GWAS is isolated to low-level, chromosomal mupR.

To study this further, we created six mupR mutants of the mupS laboratory strain SH1000 by the selection method above. The mutants were screened using the PCR-based method, and

were also sequenced to ensure that there were no other mutations. When the toxicity of these mutants were analysed, they showed decreased resistance compared to SH1000. This suggests not only that the association seen in the GWAS is true, but that this applies to more than the ST239/USA300 backgrounds. Interestingly, this SNP has been shown not to affect fitness of the strains carrying it; *ileS* being an essential gene, perhaps this is unexpected. In the absence of mupirocin, there is no difference in the growth rate of SH1000 and its mupR mutants, which is in line with the previous studies, however, in the presence of mupirocin, the mutants grow slower, needing up to 48h. In the presence of sub-inhibitory concentration of mupirocin (0.1µg/ml), one of the SH1000-mupR mutants (MY40) grows slower than SH1000, despite growing faster initially (fig. 3.9). This would suggest that exposing these strains to sub-inhibitory mupirocin, where the strains have to induce the stringent response, reveals a possible fitness burden incurred by the V588F IleRS mutation.

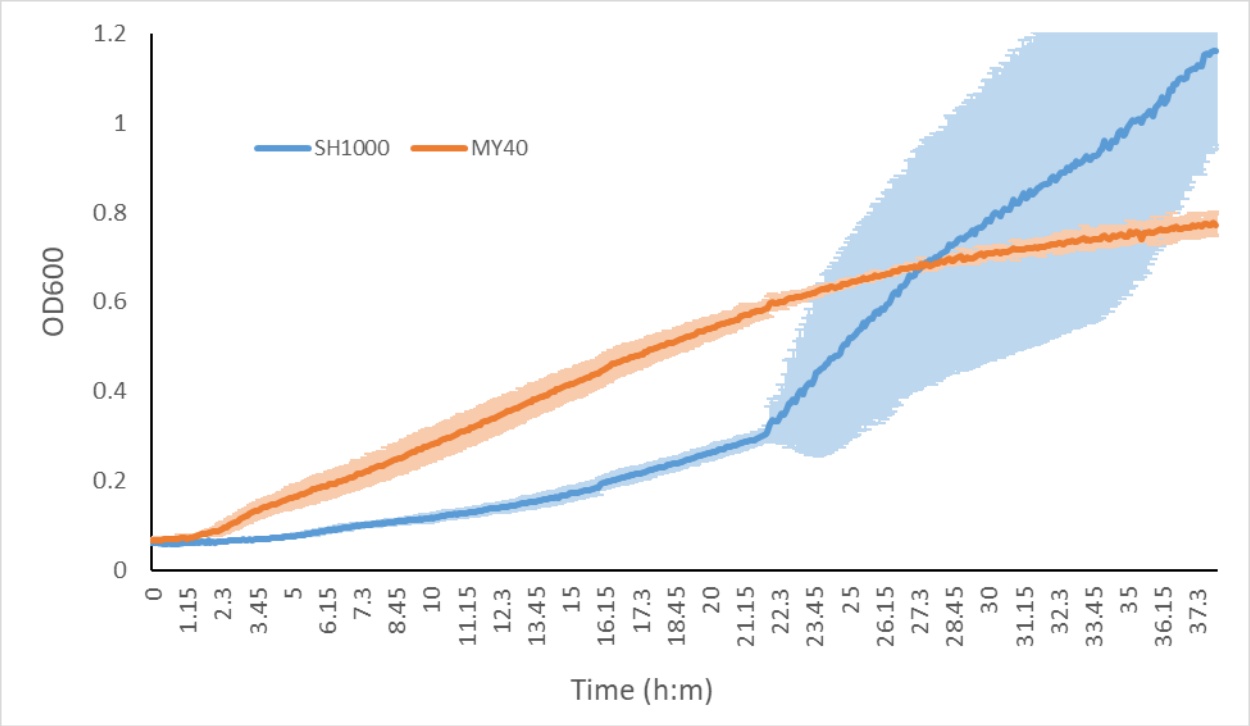


Fig. 3.9: growth of SH1000 (mupS) and MY40 (mupR) in the presence of sub-inhibitory mupirocin (0.1µg/ml) (n=3). While MY40 grows faster for ~23h, after ~37h SH1000 is growing faster than MY40, indicated by the error bars being separate at these points. The error bars showing the 95% confidence interval.

In conclusion; the G1762T SNP in *ileS*, leading to V588F mutation in the protein, was found to be associated with a change in toxicity when applying a Genome Wide Association Studies on ST239 and USA300 MRSA isolates. We first verified that the mupR SNP does indeed confer mupR, and used this to select for mutants containing the mupR SNP from mupS parents – SH1000 being the one used. We then screened SH1000-mupR mutants for toxicity, and found that it was significantly reduced (p=0.02). Previous studies had shown that the mupR SNP does not affect the fitness of strains harbouring it, and this was the case for the SH1000-mupR mutants in the absence of mupirocin. Looking at the toxicity of strains carrying *mupA* showed no difference in toxicity, therefore this phenomenon only applies to low-level mupR.

Chapter 4

- Elucidating the Mechanism Behind the Reduction in Toxicity Seen in Mupirocin Resistant *Staphylococcus aureus* Strains

Introduction

Antibiotic resistance is beneficial to the bacteria in the presence of the antibiotic, however this frequently comes at a fitness cost[250]. When bacteria acquire resistance to antibiotics, there are often compensatory mutations seen in these resistant bacteria which enables them to maintain a comparable level of fitness to antibiotic sensitive bacteria. Certain mutations are only observed in bacteria which are resistant to multiple antibiotics[250]. This would mean that antibiotic resistance becomes fitness cost-free, and thus more likely to be maintained in the population as it gives the bacteria the advantage of being antibiotic resistant without its fitness being affected. One antibiotic resistance mechanism considered to be fitness cost-free is a mutation leading to low-level resistance to mupirocin; this is a single nucleotide polymorphism (SNP) at position 1,762 in the *isoleucyl-tRNA synthetase* gene (*ileS*) of *Staphylococcus aureus*, where a G changes to a T (G1762T) (mupR SNP)[249,251]. This mupR SNP changes the amino acid at position 588 from valine to phenylalanine (V588F) in the product, IleRS, and this confers low-level resistance to mupirocin. IleRS charges tRNA with isoleucine (Ile), and mupirocin blocks this activity by binding to the active site of IleRS. Position 588 is situated in in the Rossman fold of IleRS, were the enzymatic activity of this protein occurs. Potentially, this V588F mutation could negatively affect the activity of the protein, thus we hypothesise that this mutation could lead to increased free Ile in the cytoplasm.

Numerous attempts to inactivate the *ileS* gene (e.g. Nebraska Tn library[170]) have failed, suggesting that this is an essential gene. It is also highly conserved in *S. aureus*, so it is perhaps unexpected that the mupR SNP in *ileS* was found not to have an effect on the fitness of the strains harbouring it[249]; if this mutation does have an effect on fitness, then it is unlikely to be

maintained as it would be a disadvantage in the absence of mupirocin. Supporting the previous observations that the mupR SNP is not associated with a fitness cost, growing mupR strains as a single culture without mupirocin showed no difference in growth rate. But what happens when they are grown together with other strains? Making the mupR strain compete for nutrients and oxygen with another strain may reveal a fitness cost that is not seen when the mupR strain is not forced to compete. Toxin production is an energetically expensive process which is readily switched off under stressful conditions[191], so is the reduction in toxicity that we have demonstrated in the previous chapter a mechanism by which these mupR strains are maintaining their fitness? - previous work in the lab showed that resistance to methicillin reduces virulence of MRSA, therefore this could potentially be the case for mupirocin resistance also.

Previously, we verified the association seen in the Genome Wide Association Studies (GWAS) between the mupR SNP and toxicity in *S. aureus*, where it was found that the mupR SNP leads to reduced toxin production in the strains which harbour it. Ile is a hydrophobic amino acid, and as such is expected to be more prevalent in proteins which are present in or act in hydrophobic environments. Cytolytic toxins function by burrowing into the membrane of target cells[199], and as the cell membrane is hydrophobic toxins could have a high Ile content. Also, membrane proteins may be high in Ile due to them being present in a hydrophobic environment. In order to narrow down the list of proteins high in Ile, a Monte Carlo simulation will be run[252]. These simulated Ile counts will then be compared to the actual Ile count, and a p-value assigned depending on how many simulated Ile counts are as “extreme” as the actual Ile count. The p-values will then be corrected using both Bonferroni and Benjamini-Hochberg correction for multiple comparisons as the simulation involves a million Ile counts for over 2,500 genes.

In this chapter, we endeavour to determine if the rate of tRNA charging is slower in mupR strains by seeing if there is more free Ile in the mupR cells. If this hypothesis is true, we then want to investigate if this affects proteins which are high in Ile; we will look at these proteins to see if any of these have a role in toxicity, and see if their translation is affected by being resistant to mupirocin. To see if the previous observation that methicillin resistance reduced

virulence of MRSA also applies to mupirocin resistance, we will also investigate if mupirocin resistance has an effect on the fitness of mupR strains.

Results

Mupirocin resistant cells contain more free Isoleucine compared to mupirocin sensitive cells.

To test our hypothesis that mupR SNP in *ileS* decreases the rate of isoleucyl-tRNA production, we quantified free Ile in the cytoplasm of SH1000 (mupS) and MY40 (SH1000-mupR) cells by LC-MS using a variation of the protocol by Sowell *et al*.[242]. The strains were grown for 18h then pelleted and washed before being lysed. Debris was removed by centrifugation, and soluble proteins (>3kDa) removed using a spin concentrator. These prepared lysates were then put through LC-MS. We found a small but significant increase in free Ile in the mupR cells (5.5µg/ml) compared to the mupS cells (5.2µg/ml) (one-tailed p=0.03) (fig. 4.1). This suggests that there could be a reduction in the rate at which isoleucyl-tRNA is formed.

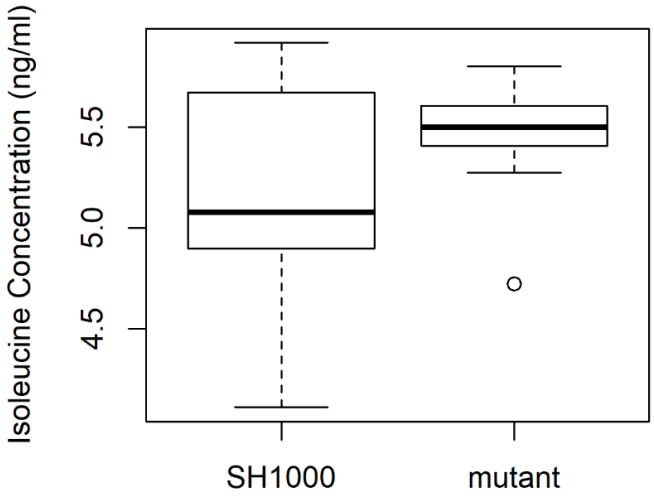


Fig. 4.1: Quantifying free isoleucine in mupS (SH1000) and mupR (MY40) cells (n=16). The cells were lysed and the lysate was analysed by LC-MS alongside known concentrations of isoleucine to determine the amount of isoleucine present. There is more free isoleucine present in mupR cells (MY40) compared to mupS cells (SH1000) (p=0.03, one-tailed T-test. n=12). The box represents the lower and upper quartiles, with the line showing the median value. The whiskers show the lower and upper extremes, and the points outside the whiskers are the outliers.

Using a Monte Carlo simulation to determine which proteins are high in isoleucine

Since there were no obvious proteins which were translated differently, we took a different approach and used a computational method to find out which proteins are significantly high in Ile. This was achieved by running a Monte Carlo simulation where each gene in two MRSA genomes (TW20 (ST239 reference strain[253]) and FPR3757 (USA300 reference strain[248])) were assigned random Ile counts (fig. 4.3). This was achieved by first finding out the total number of amino acids in the whole genome. The number of genes present in the genome was also calculated, and this was subtracted from the total amino acids count generated above to create a new total; this is to account for the non-variable first amino acid in each polypeptide. Then, a random number was generated up to that new total. For example, if the total number of amino acids was A and the number of genes was B, then the maximum random number generated is D (D = A – B). A threshold was set to the number of total Ile in the genome (C), and if the random number generated was below the threshold, it was counted as an Ile. If the random number was above the threshold, it was counted as not Ile. This was repeated for each amino acid in the gene minus one, as the first amino acid is always a methionine, which generates a simulated Ile count. This process was repeated 1 million times for each gene, then repeated for all the genes in the whole genome. The number of these simulated Ile counts which were greater-than-or-equal to (≥, E) or less-than-or-equal-to (≤, F) than the actual Ile count of the gene were determined, then the smaller of these numbers (E or F) was divided by 1 million (the number of simulated Ile counts) to determine a p-value.

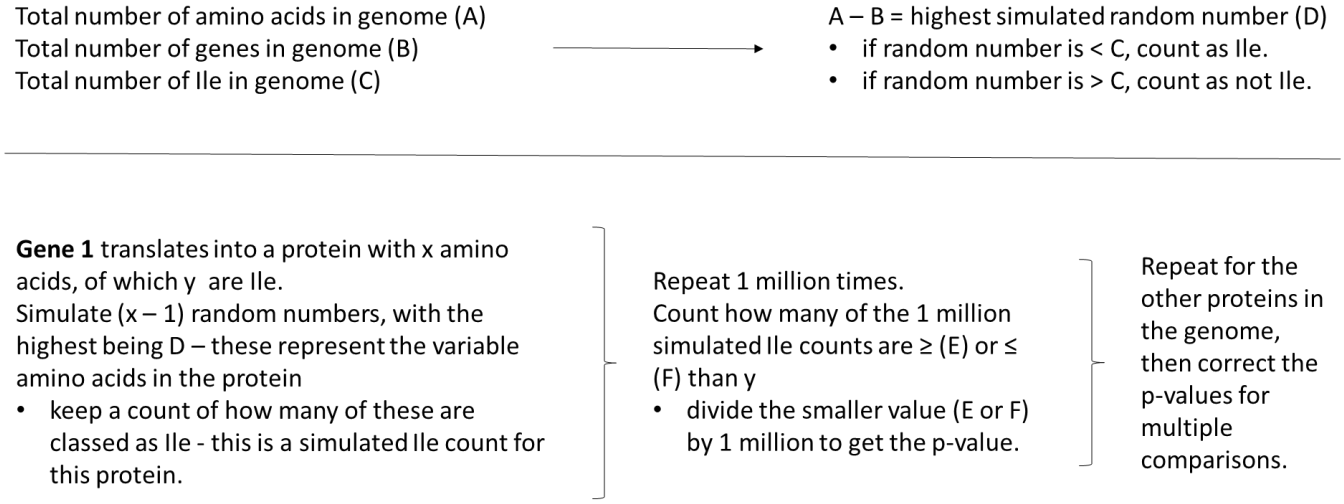


Fig. 4.3: a schematic describing the approach used for the Monte Carlo simulation of Ile count.

Due to the number of genes involved, it is necessary to use a method to correct for multiple comparisons. A p-value of 0.05 means that there is a 5% chance of a result occurring by chance, therefore if we are looking at 2,500 genes then there would be 125 genes which are significantly different in Ile only by chance. Corrections for multiple comparisons take this into account by lowering the threshold of significance. Here the stringent Bonferroni and the less stringent Benjamini-Hochberg corrections were used (Appendix A, (p107)); under the Bonferroni correction, there are 48 genes in ST239 (Appendix A, table 1 (p107)) which are significantly different in Ile, and 51 such genes in USA300 (Appendix A, table 2 (p119)). Both these include AgrC, which will be examined in the next section. However, these numbers increase to 887 genes in ST239 and 828 genes in USA300 which are significantly high in Ile when the Benjamini-Hochberg correction is used. The Benjamini-Hochberg correction includes the genes which are significant under the Bonferroni correction, so AgrC is also significant using the Benjamini-Hochberg correction. Genes which become significant under the Benjamin-Hochberg correction include AgrB, a part of the *accessory gene regulator* (*agr*) quorum sensing system, and one of the four α-Phenol Soluble Modulins (α-PMSs) in ST239, as well as 3 of 4 in USA300. PSMs are small peptide toxins, and the three significant α-PSMs have an Ile content of >25%. This is extremely high considering that the genome average is ~9%, but due to their small size these are only significant under the Benjamini-Hochberg correction. The fourth α-PSM (PSM-α3) has no Ile, and this is not significant under either correction method.

Using an alternative method to identify proteins significantly high in isoleucine

To back up the simulation carried out above, Dr Mario Recker created a regression model of the Ile counts in the ST239 and USA300 genomes (fig. 4.4). Here, the purple line represents the average Ile count, and the dark grey area represents the 95% prediction interval; 95% of the Ile counts are expected to fall within this area given the frequency of Ile in the genome. The Ile counts were transformed using a Box-Cox power transformation to limit variability in the data, then plotted onto the graph. AgrC, AgrB and the α-PMSs all fall outside the 95% prediction interval, so these are all significantly high in Ile – this further supports the findings of the simulation run above.

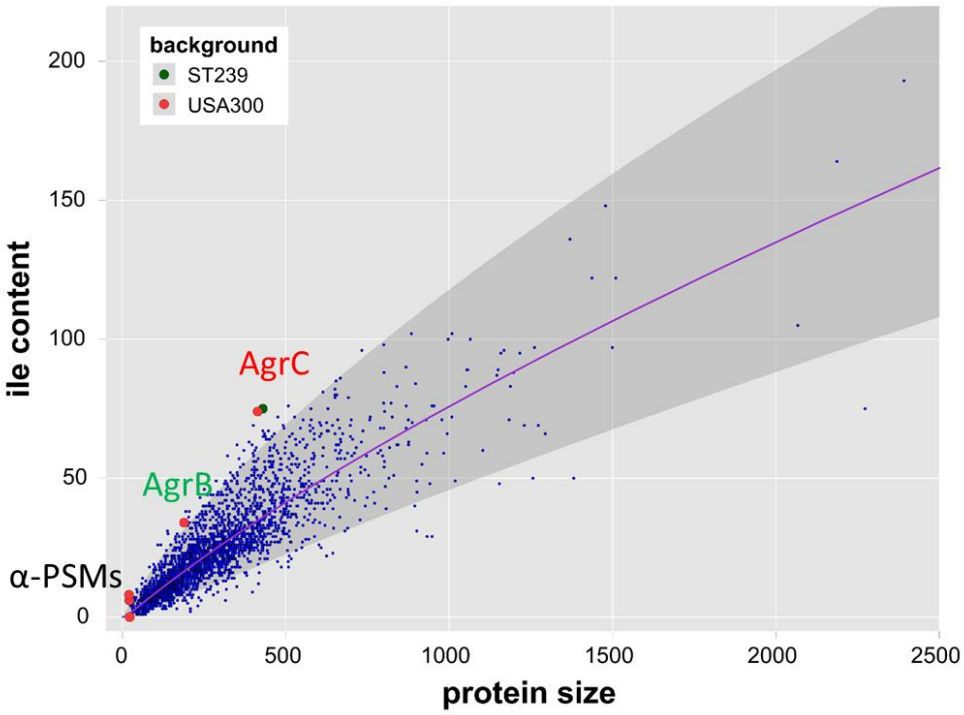


Fig. 4.4: a regression model showing the distribution of Ile counts in the proteins from ST239 and USA300. When fitting the Ile counts onto a regression model, AgrC, AgrB and the α-PSMs all fall outside the 95% prediction interval (dark grey area), indicating that they have a significantly high Ile content. The Ile counts for all genes were transformed using a Box-Cox power transformation prior to plotting. The purple line represents the average Ile count, with the dark grey area representing the 95% prediction interval – this is where 95% of Ile counts are expected to be, given tis frequency. Data courtesy of Dr Mario Recker.

AgrC, a protein high in isoleucine, is translated slower in a mupirocin resistant strain

From the Monte Carlo simulation, we discovered that AgrB and AgrC were significantly high in Ile content (Benjamini-Hochberg corrected for AgrB, Bonferroni and Benjamini-Hochberg corrected for AgrC), which was verified in the regression model. The *accessory gene regulator* (*agr*) system is a global regulatory system in *S. aureus* which controls many genes in response to population density, and is an auto-amplification loop activated at high population density[225]. Two of the gene of this system, *agrB* and *agrC*, code for membrane proteins; AgrB, which processes and exports the signal peptide AIP, and the sensor histidine kinase AgrC. Toxins are one of the cell processes up-regulated by the *agr* system.

We next wanted to study the effect of mupR on AgrC translation and transcription, and to achieve this we created isogenic mupS and mupR strains which were not able to produce AIP but were able to respond to exogenous AIP; into *agr* knock-out mupS (MY18) and mupR (MY41) strains, a plasmid containing the *agr* P3 promoter, *agrC(his)* and *agrA* (pAgrC(his)A (from collaborators in Nottingham)) was transformed. These "*agr* replaced" strains ensure that the *agr* system is activated to the same extent and the addition of the hexa-his tag enabled visualisation of the AgrC protein (fig. 4.5). This removes the variability in general *agr* activity between strains which could occur naturally, to give a comparable result between the mupS and mupR strains.

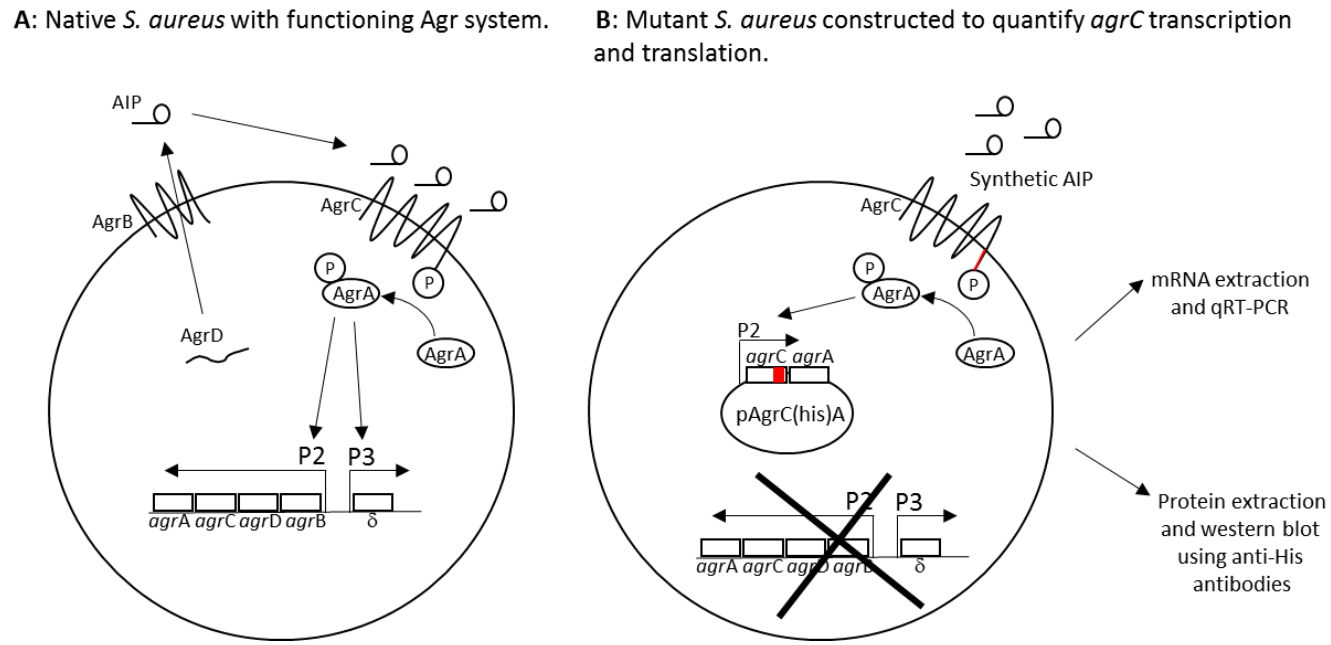
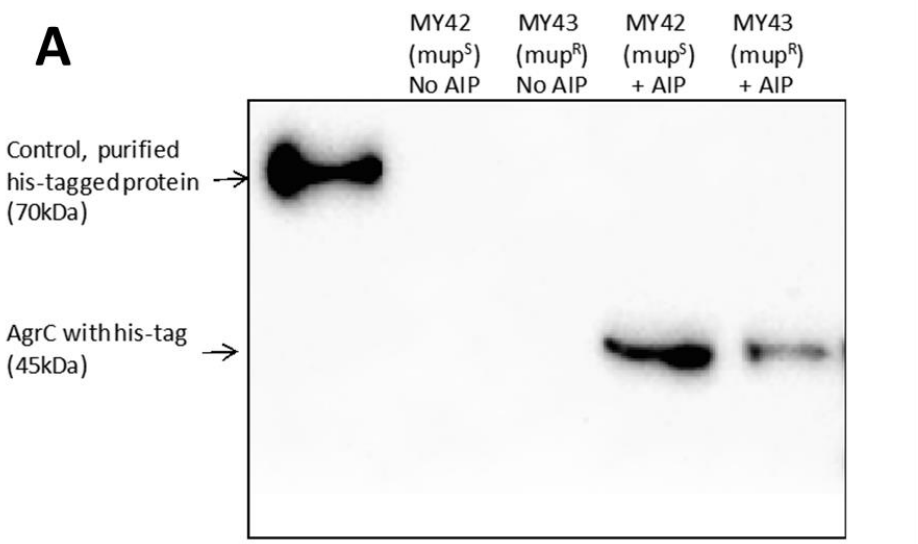


Fig. 4.5: construction of the "*agr*-replaced" strain to facilitate the analysis of *agrC* protein and mRNA. This become an *agr* knock-out strain with the ability to respond to exogenous AIP. The native *agr* system was replaced with a P3-lux and erythromycin resistance cassette, then a plasmid containing AgrC(his)A was transformed into the knock-out strain. This allows the strain to respond to synthetic AIP in the media, without having the ability to produce AIP. The his-tag on AgrC allows for visualisation of this protein by Western Blotting.

Translation of AgrC was quantified by Western blotting; MY42 (mupS) and MY43 (mupR) were induced with synthetic exogenous AIP for 1-3h, and then were lysed. The proteins in the lysates were then separated on SDS-PAGE gels and then transferred onto nitrocellulose

membrane for the Western blot. At 1h post-induction, there was approximately 2x more AgrC(his) in MY42 compared to its MY43 using densiotometry (fig. 4.6 A). This was not seen at 2h and 3h post-induction. So, it seems that mutating *ileS* slows the translation of AgrC early on, with the mupR strain producing the same amount of AgrC as the mupS strain later on.

A



B

mupS (*agrC* Ct/*gyrB* Ct)	mupR (*agrC* Ct/*gyrB* Ct)	ratio mupR/mupS
0.95 ± 0.05	0.98 ± 0.04	1.04

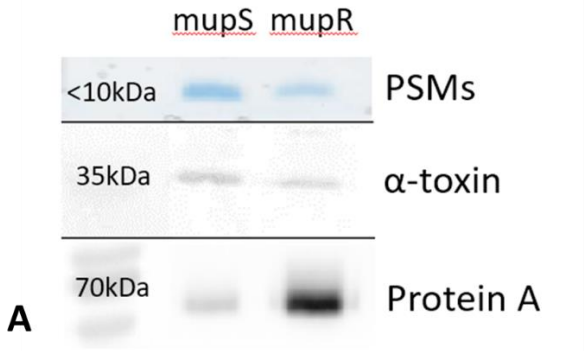
Fig. 4.6: Western blot analysis of AgrC(his) produced by mupS and mupR agr-replaced strains (n=3). Both mupS and mupR versions of the *agr*-replaced strain were grown overnight, then diluted 1:500 into fresh media, then grown for 2h before adding AIP. 1h after this addition of AIP, there is less AgrC(his) in the mupR strain compared to the mupS strain, as shown by a reduced intensity in the band (A). However, this is not reflected in the transcription (B); the Ct threshold values from the qPCR run for *agrC* were converted into ratios by dividing it with the Ct value of a housekeeping gene, *gyrB*, which is constantly expressed. The values are average of three repeats, shown with the 95% confidence interval.

As the translation of AgrC is affected in MY43, we next sought to see if this was due to differences in transcription. Transcription of *agrC* was tested by quantitative reverse-transcription PCR (qRT-PCR); first, the primers were tested by a standard PCR and genomic *S. aureus* DNA to ensure that they were specific to *agrC*. Once verified, qRT-PCR was performed on cDNA samples – these samples were taken from MY42 and MY43 1h post-induction, as this

was where the difference in translation was seen. The qRT-PCR showed that there is no difference in the rate of *agrC* transcription (fig. 4.6 B), therefore the difference seen in the quantity of AgrC(his) is due to a reduced translation rather than transcription. Therefore, this suggests that mutating *ileS* to become mupR affects the rate of IleRS activity, which in turn affects AgrC(his) translation.

Does the temporal slowing of AgrC translation affect downstream processes regulated by the agr *system?*

As early AgrC translation is affected in mupR strains, we sought to determine if this had a knock-on effect on genes regulated by *agr*. Therefore, we looked at the amount of Protein A and α-toxin produced by SH1000 and MY40 by Western blotting; Protein A is down-regulated by the *agr* system, while α-toxin is up-regulated[223]. We also looked at the production of Phenol Soluble Modulins (PSMs), as three of the four α-PSMs were shown to be high in Ile and they also regulated by the *agr* system[181]. It was clearly visible from the SDS-PAGE that there is more α-toxin and PSMs produced by SH1000 compared to MY40, while more Protein A was produced by MY40 (fig. 4.7). This suggests the possibility of the *agr* system being affected by the delayed translation of AgrC.



	SH1000 (mupS)	MY40 (mupR)	p-value
PSMs	3057.92	2246.93	0.0004
α-toxin	6493.92	3817.46	0.02
protein A	1216.34	16535.88	<0.0001

B

Fig. 4.7: analysis of proteins controlled by the *agr* system to see if the early reduction in translation of AgrC has an effect on these proteins; part A shows the gel/Western blot and part B shows the densiotometry results using imageJ as a measure of quantification (n=9 for PSMs, n=6 for α-toxin and Protein A). p-values from T-tests. Phenol Soluble Modulins (PSMs) are short peptide toxins, and along with α-toxin is upregulated by the *agr* system. Also, three of the four α-PSMs have an extremely high isoleucine content, however due to their short nature this is not significant using the Monte Carlo simulation. Both PSMs and α-toxin show reduced production in the mupR strain. Protein A is an immune evasion protein which binds to antibodies, and this is down-regulated by the *agr* system. There is a higher amount of Protein A produced by the mupR strain.

Mupirocin resistance in an agr-ve *strain affects its fitness*

The above data suggests that the reduction in toxicity seen in mupR strains could be through the delay in *agr* activity in the mupR strains. The initial delay in AgrC translation seemingly is sufficient to cause this slight delay in *agr* activity, however other mechanisms may also be reducing the toxicity of mupR strains. As mentioned before, toxin production is an energetically expensive process which is switched off readily if the growth conditions become sub-optimal. Also, a previous study in our lab showed that methicillin resistance reduced toxin production in MRSA, offsetting some of the associated fitness costs.

Knowing that this mutation has been reported to have no effect on fitness despite being an essential gene, but having detected that they have lowered their toxin production, we sought to understand the fitness consequence of mupirocin resistance in greater detail. To achieve this we performed some relative fitness competition assays; by growing mupS and mupR together, they are forced to compete for nutrients and oxygen, thus effects on fitness not

seen in a pure culture may be exposed here. To investigate the role of toxin production in the maintenance of competitive fitness, we utilised four strains; mupS and mupR strains with an intact *agr* system (which can produce toxins) and mupS and mupR strains with the *agr* system knocked-out (which do not produce toxins). The four strains (SH1000 & MY40 (*agr*+ve), and MY18 & MY41 (*agr*-ve)) were grown overnight individually, and then competed against another strain by culturing together, as "SH1000 and MY40" (SH1000-pair) and "MY18 and MY41" (MY18-pair).

The overnight cultures of the four strains were diluted, then equal volume of the dilute mupS and mupR strains were pipette into fresh media. One culture contained the SH1000 (Agr+ve)-pair, while the other contained the MY18 (Agr-ve)-pair. These were grown for 24h, diluted then the plated onto fresh agar plates – one containing mupirocin and the other containing no antibiotic. This enabled us to obtain a colony forming unit/ml (CFU/ml) count for each strain. By subtracting the number of colonies on the mupirocin plates from the number of colonies from the no antibiotic plate, it is possible to obtain the count for the mupS strain. These CFU/ml values were converted into Malthusian parameters using the equation in the materials and methods to show the maximal exponential growth of each strain, and compared as the SH1000-pair and MY18-pair. The SH1000-pair showed no significant difference in fitness ($p=0.6$), but MY41 showed reduced fitness when grown together with MY18 ($p=0.02$) (table 4.1). The data show that MY41 (the *agr*-ve mupR strain) has a significantly lower MP compared to its mupS counterpart, MY18, which shows that MY41 does not grow as fast as MY18. Therefore, the competitive fitness of MY41 is lower than that of MY18. The *agr*+ve mupR strain MY40 does not show this reduction in MP, thus does not show a reduction in competitive fitness. The difference between MY41 and MY40 is that MY40 produces toxins while MY41 does not, and this means that toxin production can be modulated in MY40 if necessary. Therefore, *S. aureus* is able to off-set the fitness cost associated with mupR mutation by reducing the amount of toxins it produces, effectively providing a cost-free antibiotic resistance mechanism.

	agr +ve		agr -ve	
	mupS	mupR	mupS	mupR
Average Malthusian Parameter	14.92425	14.75803	15.00947	14.37296
t-test	0.603316		0.01982	

Table 4.1: analysing the fitness of mupS and mupR strains by growing them in a co-culture (n=19 for agr+ve, n=18 for agr-ve). When SH1000 (mupS) and MY40 (mupR) have an intact *agr* system, and there is no difference in fitness between these strains (p=0.6). However, when the *agr* system is knocked-out (MY18 (mupS) and MY41 (mupR)), fitness is reduced in the mupR strain (p=0.02, T-test).

Discussion

Here, we set out to determine the mechanism behind the change in toxicity seen in low-level mupirocin resistant (mupR) strains carrying the V588F mutation. The product of the *ileS* gene charges tRNA with isoleucine (Ile), and with the amino acid 588 predicted to be in the Rossman fold where this activity occurs, we hypothesised that mutation could slow the rate of tRNA charging. If this were the case, we would expect more free Ile in the cytoplasm of mupR strains compared to mupirocin sensitive (mupS) strains. LC-MS analysis showed an increased amount of free Ile in mupR cells, so this suggests that the rate at which isoleucyl-tRNA forms is slower in mupR cells. This would then possibly affect proteins with high Ile content.

Due to decreased isoleucyl-tRNA in mupR cells, proteins high in Ile may be translated slower in mupR strains compared to the mupS strains. Therefore the next step we took was to establish which proteins are high in Ile; this was achieved by simulating Ile counts for each protein, and comparing this to the actual Ile count. Due to the high number of simulated counts generated for each gene, it was necessary to correct for multiple comparisons. Here, we used both the stringent Bonferroni method and the less stringent Benjamini-Hochberg method. This then gives a list of proteins which are significantly high in Ile, which is different to those which have a high percentage of Ile; if the gene yielded a short peptide, it is statistically more likely to simulate an Ile count which is extremely high compared to genes which produce long peptides. For example, three of the four αPSMs have an Ile count of >25%, which is very high compared to the genome average of 9%. These are short peptide toxins which form pores in the target cell membrane. Due to their very short peptide length the

chances of a simulated Ile count being that high is high, therefore this did not reach statistical significance under Bonferroni correction.

One of the proteins which was highly statistically significant after Bonferroni correction, both from the ST239 and USA300 genomes, was AgrC. This is a component of the *accessory gene regulator* (*agr*) quorum sensing system. AgrC is a membrane bound histidine kinase which phosphorylates the response regulator (AgrA) when the signal (autoinducing peptide (AIP)) binds. In order to see if translation of AgrC was slower in mupR strains, we devised a system whereby the activation of the *agr* system was controlled. This ensures that the system is activated to the same degree, so differences in overall *agr* activity between strains will not affect the results. A hexa-his tag was added to the AgrC by collaborators to allow the detection of AgrC; rather than having an antibody or probe against AgrC, this enables us to use one against the his-tag. Western blots showed a fainter AgrC band from mupR lysates taken 1h after induction, but this was not maintained at later time points (2-3h). This shows that at the early stages, translation of AgrC is slower in mupR strains. However, qRT-PCR on cDNA from cells 1h post-induction showed no difference, therefore the decrease in AgrC protein is not due to a decrease in *agrC* transcription. As this was one of the proteins identified as being high in Ile, it would seem that the reduced isoleucyl-tRNA present in mupR cells is causing AgrC to be translated slower temporarily. This difference is made up later on, therefore there is a possibility that overall *agr* activity is not greatly affected. However, as the difference in toxicity between mupS and mupR strains is small, the early delay of AgrC translation in mupR strains may be sufficient to cause this difference.

Protein A and α-toxin are both regulated by the *agr* system[223], and as such if there was a difference in *agr* activity then there would be a difference in expression of these proteins. A group of toxins are the phenol soluble modulins (PSMs), and as mentioned above three of the four α-PSMs have a high Ile content. When Protein A, α-toxin and PSMs were examined by Western blots and SDS-PAGE, there was a difference seen between mupS and mupR; there was more Protein A in the mupR strain, and more α-toxin and PSMs in mupS strains (all $p=<0.05$). This indicates that perhaps there is a slight delay in *agr* activation.

When the strains are grown individually without mupirocin, there is no difference in growth rate between mupS and mupR strains, including strains which have the *agr* system knocked-out; these strains cannot make toxins due to the lack of *agr* system. However, when we compete mupS and mupR strains in a co-culture, we start seeing a difference in fitness – the SH1000-pair, with an intact *agr* system, are equally fit, as shown by their similar Malthusian Parameter (MP), meaning they are able to grow to the same extent. The mupR strain of the MY18-pair, MY41, with the *agr* system knocked-out, shows reduced fitness compared to its mupS equivalent (MY18), as shown by its reduced MP, meaning MY41 does not grow as well as MY18 in a co-culture. When these strains are cultured together, the two are having to compete for nutrients and oxygen, therefore this is more likely to reveal reduction in fitness compared to growing the strains in a pure culture. This suggests that having the ability to produce toxins means that this can be reduced to alleviate any fitness costs associated with becoming mupR, but taking this ability away makes the mupR strain unable to retain fitness in a competition. This is an elegant example of how bacteria are able to adapt to stressful situations in order to survive and thrive.

As we discovered in the previous chapter, the V588F IleRS mutation leading to low-level mupR causes a reduction in toxicity. Here, we found evidence to suggest that this mutation lowers the rate of IleRS activity, in that there is more free Ile found in mupR cells. One of the proteins which we found to be high in Ile, AgrC, is translated slower in mupR strains during the early phases of its translation. Toxin production is an energetically expensive process, and we found that losing this ability made the mupR strain less fit compared to its mupS counterpart, while this was not seen when the strain had the ability to produce toxins. This suggests that the combination of slower AgrC translation and the need to maintain fitness in the face of IleRS mutation leads to the reduction in toxicity seen in mupR strains.

Future directions

In this research, we saw a slight increase in the amount of free Ile present in mupR cells which would indicate that there is less isoleucyl-tRNA present in these cells. This could be investigated further, by quantifying the amount of isoleucyl-tRNA directly or by comparing the activity of the mupS (V588) and mupR (F588) IleRS. Here, we did not quantify the level of IleRS

present in the mupS and mupR cells, so this is something that could be investigated. If there is more IleRS present in mupR cells, then this may counteract any delay in its activity.

This research also revealed a decrease in AgrC translation at the early stages of its induction. This protein was found to be significantly high in Ile, therefore is this delayed translation seen in all proteins high in Ile? Or is there a trade-off where essential proteins high in Ile are preferentially translated, so that there is no delay in the translation of these proteins? These are questions that could be investigated in the future.

Chapter 5

- Finding Novel Genes Affecting Lipase Activity of *Staphylococcus aureus*

Introduction

Lipases are enzymes which hydrolyse ester bonds in lipids, and in *Staphylococcus aureus* these enzymes are involved in lipid metabolism as well as virulence[219]. Part of the defence mechanisms used by humans is the production of fatty acids on the skin, and the most effective of these against staphylococci is *cis*-6-hexadecenoic acid, also called sapienic acid (C16:1 Δ6)[220]. It has been noted that individuals who are prone to atopic dermatitis produce reduced amounts of hexadecenoic acid, which is linked to increased *S. aureus* carriage in these individuals[254]. Also, lipases can hydrolyse these lipids into free fatty acids, which are also toxic to staphylococci. A study observed an increased presence of bactericidal fatty acids in staphylococcal abscesses, which seems to be a mechanism used by the host immune system to control and eliminate staphylococci[255]. The authors found that these fatty acids were long-chain fatty acids released from glycerides, most likely by the action of staphylococcal lipases. Therefore, it seems paradoxical that staphylococci produce lipases, which would cause the formation of compounds which are toxic to staphylococci. Another study showed that lipase from USA300 strains were capable of releasing growth-inhibitory free fatty acids and a study by White *et* al. showed that phospholipase C from USA300, which is specific for phosphatidylinositol, aided survival in blood[256].

A study by Hu *et al*. found that biofilm formation was decreased when a lipase gene was knocked-out, and that inoculation with the lipase mutant led to reduced peritoneal abscesses and lower bacterial loads in the organs of inoculated mice[221]. This supports findings from other studies which found *S. aureus* strains derived from deep infections produced more lipase compared to those isolated from superficial infections[257]. This indicates that lipase also has a role in invasion, and also in degradation of host tissue for nutrient acquisition. Studies investigating the link between staphylococcal lipase and immune responses have found that

while low concentration of lipase was chemotactic for granulocytes, higher concentrations seemed to immobilise granulocytes[258]. Also, the same study found that pre-incubation of granulocytes with staphylococcal lipase reduced its ability to kill bacteria by phagocytosis, mainly due to reduced phagocytic uptake.

Lipase production is regulated by the *accessory gene regulator* quorum sensing system; at high cell density, the *agr* system is activated, and this up-regulates lipase production[225]. Mutants lacking the *agr* system shows reduced lipase production, and this could contribute to reduced virulence of these mutants in a mouse model. Also, *sarA* mutants showed a reduction in lipase production, suggesting that the *sar* regulatory system also plays a part in regulating lipase production[259].

Staphylococcal lipases are also used in biotechnology as catalysts; their reactions are highly specific, and they are not only able to hydrolyse a wide range of substrates but can also synthesise fatty acids[219]. Using enzymatic catalysis therefore is more favourable than using chemical reactions because of its specificity, which in turn reduces unwanted reactions and prevents toxic by-products from forming[260]. These properties then make downstream separation simpler and as less energy is required for these processes the use of lipase more cost-effective compared to chemical reactions. Lipases are used in many industries including food and pharmaceutical industries, and also in the production of biodiesel[222].

In this project, we wanted to identify novel regulators of staphylococcal lipase activity. To achieve this, we screened a sequenced collection of Methicillin-Resistant *S. aureus* (MRSA) isolates for variation in lipase activity, and then compared these results with the genome sequences in a Genome Wide Association Study (GWAS) to see what genomic polymorphisms were associated with the variation in lipase activity[239]. As seen in the previous chapters and from other research in our lab, GWAS have given us an insight into novel ways which virulence, such as toxin production, is regulated, therefore we sought to achieve the same with lipase production.

Results

Developing a high-throughput method for analysing lipase activity

Assays to quantify lipase activity are commonly carried out in a liquid buffer system. A method developed by Gupta *et al*. used a Tris-based buffer with gum Arabic with the substrates *para*-nitrophenol butyrate (pNPB, short chain) and *para*-nitrophenol palmitate (pNPP, long chain)[261]. The substrates are suspended in isopropanol, and are then mixed in a 1:9 ratio with the Tris-based buffer. The study by Gupta *et al*. found that the fatty acids released from the substrates by the lipase activity dissolved into the buffer with the addition of Triton-X100, which prevented the assay from becoming turbid. The assay by Gupta *et al*. was carried out as 1ml reactions using cuvettes and a spectrophotometer, but we reduced the total volume to 200µl so that the assay can be carried out in a 96-well microtitre plate. This makes it possible to assay multiple strains simultaneously, making the assay more high-throughput and facilitating the analysis of a whole strain collection.

A standard curve (fig 5.1) was prepared by dissolving known concentrations of *para*-nitrophenol (pNP), the yellow compound conjugated to the fatty acids in the substrate which is released by lipase activity. The absorbance was read in a plate reader every 5 mins for 1h, which gives a kinetic curve from which µM *para*-nitrophenol (pNP) released/min can be determined; this was achieved by finding out the peak amount of pNP release and dividing it by the number of minutes it took to reach this peak.

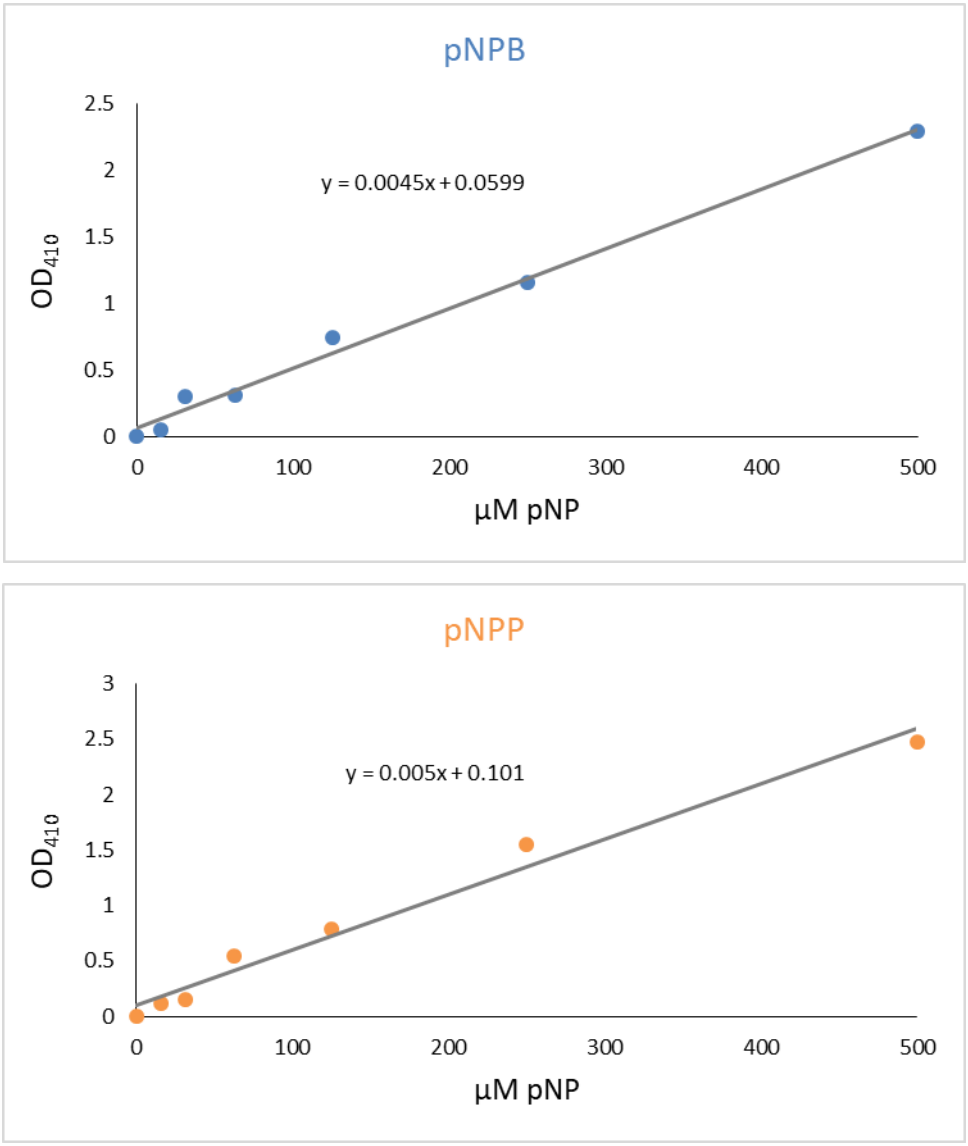


Fig. 5.1: example standard curves used for the lipase assay (n=2). Known concentrations of *para*-nitrophenol (pNP) was dissolved into the assay buffer, and their OD_{410} was read. This example is of the pNP dissolved into the *para*-nitrophenol butyrate (pNPB) assay buffer. The equation shown was generated by Excel from the data, and this was used to calculate the pNP concentration of the samples from their OD_{410} readings.

Lipase is secreted, thus released into the supernatant therefore the cultures were grown for 18h then pelleted to collect the supernatants. In the preliminary assays, an *agr*+ve strain (SH1000) and an *agr*-ve strain (MY18) were used, as the *agr* system is known to regulate lipase activity. These strains gave a consistent difference in lipase activity (fig. 5.2), therefore SH1000 and MY18 were included in all further assays to ensure that it was functioning as expected.

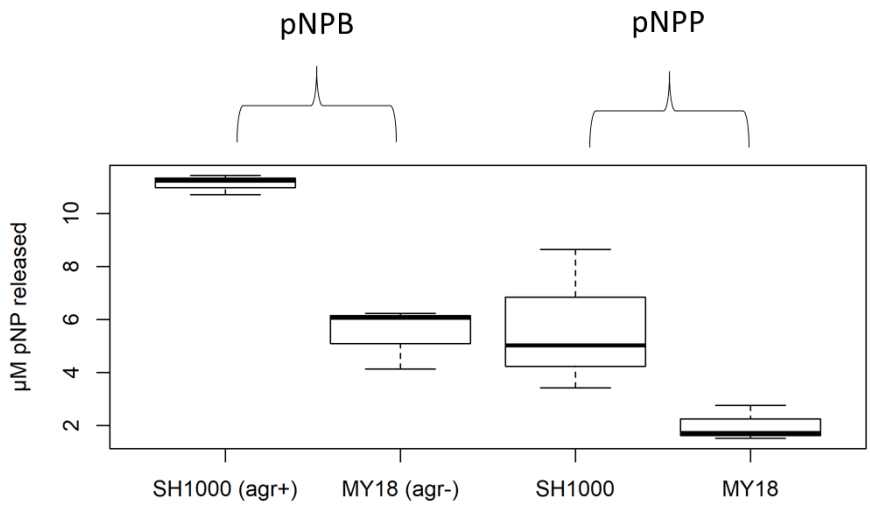


Fig. 5.2: lipase activity of an *agr*+ve strain, SH1000, and its *agr*-ve strain, MY18, shows that the activity of the *agr* system is involved in lipase production (n=3). As these strains consistently showed a difference in lipase activity, these were chosen as the control strains for subsequent lipase assays. pNPB (*para*-nitrophenol butyrate) is the short-chain substrate and pNPP (*para*-nitrophenol palmitate) is the long-chain substrate. The box represents the lower and upper quartiles, with the line showing the median value. The whiskers show the lower and upper extremes, and the points outside the whiskers are the outliers.

Identifying novel genes associated with lipase activity in ST239 strains

To identify genes which affect lipase production, a Genome Wide Association Study (GWAS) was carried out by Dr Mario Recker using 95 ST239 MRSA strains; the ST239 lineage of MRSA is healthcare-associated, and is a global clone with prevalence in Asia, South America and Eastern Europe[253]. The results showed that in general, these ST239 strains produce a low level of lipase but there were differences between individual isolates (fig. 5.3). It was also noticed that activity towards the short-chain substrate (pNPB) was greater than towards the long-chain substrate (pNPP). Dr Recker then performed a GWAS using PLINK[262] with this data and the genome sequence of these strains, and this approach identified various SNPs which were associated with a change in lipase activity; this included a lipase gene as well as *agrC* (table 5.1).

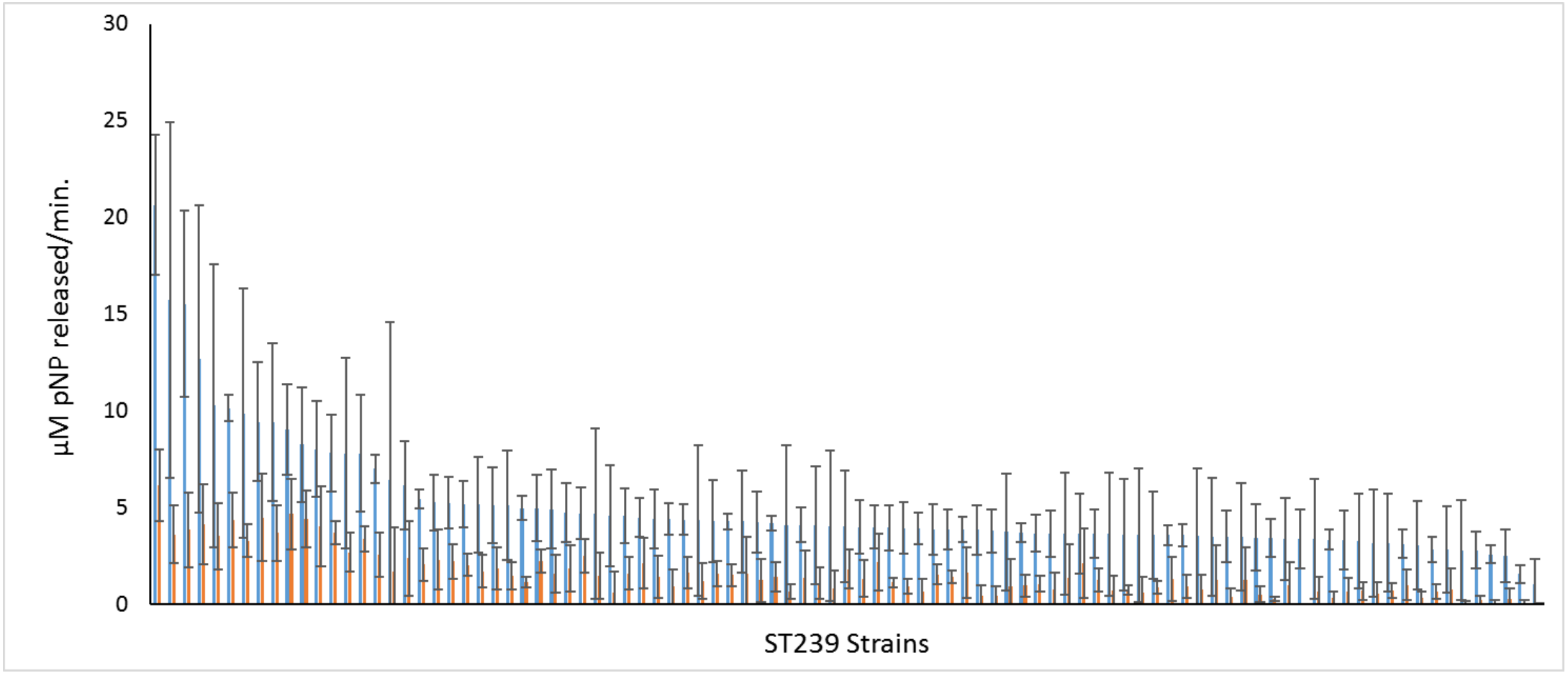


Fig. 5.3: a graph showing the lipase activity of ST239 strains, represented by the amount of para-nitrohenol (pNP) released per min (n=3). The blue bars show the release from a short chain substrate, para-nitrophenol butyrate, and the orange bars show the release from a long chain substrate (para-nitrophenol palmitate), with the 95% confidence interval shown as error bars. From the graph, it is visible that there is variation in lipase activity between the ST239 strains, with higher activity seen for the short chain substrate.

Table 5.1: The genomic polymorphisms which showed the most significant association with lipase activity as identified by the GWAS. When comparing strains with and without the below SNPs, there was a statistical difference in lipase activity. In some cases, there were multiple SNPs in the same gene or intergenic location, and these are only listed once in the table; noted by "From..." and "x..." in the left two columns. These hits have a $-Log_{10}$ p-value of >5, which is the equivalent to p<0.00001. Some of the hits are part of a prophage which had integrated into the genome (fig. 5.4), in the "*geh*" and "*hlb*" genes, a lipase or a phospholipase precursor gene.

position in genome	**-LOG10(p)**	**gene locus/name**	**description**	**NE mutant**
From 419255	11.93632147	intergenic *lytA*/SATW20_03890 x3	*lytA* – autolysin Part of a prophage in lipase precursor gene/*geh*	
2147902	9.955025417	SATW20_19910 (phage antirepressor protein)	Part of a prophage in phospholipase C precursor gene/*hlb*	NE1786
2942464	9.357797992	*clfB*	Clumping factor B	NE391
From 2147007	7.79207275	SATW20_19900 (phage protein) x2	Part of a prophage in phospholipase C precursor gene/*hlb*	(no sig. match in USA300)
2144787	7.270977981	SATW20_19830 (phage protein)	Part of a prophage in phospholipase C precursor gene/*hlb*	
From 390115	6.913652874	SATW20_03490 (phage protein) x2	Part of a prophage in lipase precursor gene/*geh*	(no sig. match in USA300)
152937	6.80695886	SATW20_01390		NE585
301089	6.80695886	intergenic *tarK*/*tarF*	*tarK* – putative CDP-glycerol: poly(glycerolphosphate) glycerophosphotransferase *tarF* – putative teichoic acid biosynthesis protein	
967656	6.80695886	SATW20_09130		NE499
972267	6.80695886	SATW20_09170		
1557020	6.80695886	*ebh*	Very large surface anchored protein	NE1
1937314	6.80695886	SATW20_17780		
2174068	6.80695886	*agrC*	Histidine kinase of the *accessory gene regulator* system	NE873
2377659	6.80695886	*atpH*	Putative ATP synthase δ chain	NE1889
2393426	6.80695886	*murA2*	Putative UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase	NE939

2608746	6.80695886	SATW20_24680		
2997008	6.80695886	*lip*	Lipase precursor	NE338
78396	6.80695886	*mecA*	Penicillin binding protein 2'	NE1868
961636	6.651881751	intergenic SATW20_09020/SATW20_09030		
From 2145704	6.41023554	SATW20_19860 x2		
From 2138296	6.403156921	SATW20_19730 x4		NE1826
2145901	6.403156921	SATW20_19870		NE1298
From 2147150	6.403156921	SATW20_19900 x3	Part of a prophage in phospholipase C precursor gene/*hlb*	(no sig. match in USA300)
From 2151200	6.403156921	SATW20_19980 x5	Part of a prophage in phospholipase C precursor gene/*hlb*	
From 2152025	6.403156921	*int* x8	*int* – integrase. Part of a prophage in phospholipase C precursor gene/*hlb*	NE327
2153030	6.403156921	intergenic *int*/SATW20_20010	Part of a prophage in phospholipase C precursor gene/*hlb*	
From 961672	6.315826777	intergenic SATW20_09020/SATW20_09030 x4		
2128192	6.218343607	SATW20_19530		NE181
1121452	6.002470016	intergenic SATW20_10520/*fmt*	*fmt* – autolysis and methicillin resistant-related protein	
From 2145746	5.850814809	SATW20_19860 x2		
2138224	5.736809249	SATW20_19730		
From 962123	5.43454713	SATW20_09040 x16		(no sig. match in USA300)
From 962662	5.43454713	intergenic SATW20_09040/SATW20_09050 x3		
From 962817	5.43454713	SATW20_09050 x4		(no sig. match in USA300)
411798	5.072590458	SATW20_03800	Part of a prophage in lipase precursor gene/*geh*	(no sig. match in USA300)
412877	5.072294037	SATW20_03810	Part of a prophage in lipase precursor gene/*geh*	(no sig. match in USA300)

Interestingly, some of the hits of the GWAS were from prophages which had integrated into either the *geh* gene or the *hlb* gene (fig. 5.4). The *geh* gene is a lipase precursor, and the *hlb* gene is a phospholipase C precursor. Potentially, the phages could have inserted into slightly different position along the *geh* and *hlb* genes in the isolates. In turn, this could mean that there is transcription of a fragment from the *geh* and *hlb* in these isolates, which may be functionally active. If this is the case, then it would be expected that lipase activity would be higher in these isolates, and these may have coincided with particular variations seen in the prophage DNA. These would then be identified in the GWAS as loci affecting lipase activity. Also, looking through the prophage genes within *geh* and *hlb* revealed that there are genes designated "*geh*" within the *geh* prophage and "*hlb*" within the *hlb* prophage. Potentially, these could play a role in lipase activity, and perhaps the polymorphisms in the prophage DNA could affect the rate at which these genes are transcribed, translated or affect any post-translational modification.

Fig. 5.4: a diagram showing how a prophage (blue box) disrupts an active gene (line). Without the insertion of the prophage, the gene is not disrupted, therefore expressed (A). However, the insertion of the prophage disrupts the gene, i.e. it splits the gene into two fragments (B). In the lipase GWAS hits, some polymorphisms were found in the prophages which had inserted into the *geh* and *hlb* genes. *geh* is a lipase precursor and *hlb* is a phospholipase C precursor.

Confirming the GWAS result – knocking out the genes of possible interest

As GWAS can produce false-positives, it is necessary to determine if these associations are in fact true. If the polymorphism had occurred in a gene, the gene can be inactivated (knocked-out) if is a non-essential gene – this is only possible if the candidate genes are non-essential, as knocking-out essential genes makes the mutant unviable. This can then be compared to a

strain with the active gene to see if the inactivation has an effect on lipase activity. Methods for inactivating genes include using transposons, such as the one used to create the Nebraska Transposon (Tn) mutant library[170]. If, for example, a SNP in the *agrC* gene was associated with a change in lipase activity according to the GWAS, then a Tn mutant from the Nebraska Tn mutant library which has a Tn in the *agrC* gene will be screened. The Nebraska Tn mutant library is a collection of strains based on a USA300 MRSA strain, JE2, where all non-essential genes are knocked-out; the mutants will be referred to as "NE mutants" from here on.

As this library is based on a USA300 strain, which is different to the ST239 MRSA strains we had analysed, it was necessary to use Basic Local Alignment Search Tool (BLAST) to find the GWAS candidate genes in the USA300 reference genome; BLAST searches for similar DNA or protein sequence, so here we use the DNA sequence of the ST239 gene of interest, and look for a match in the USA300 sequence. For example, *murA2* was identified in the GWAS as being associated with lipase activity, then BLAST was used to identify the region in the USA300 reference genome which matched *murA2*, which was *murA*. From there, we looked through the NE mutant library to see if *murA* had been knocked-out. We displayed the GWAS data to show the most significant associations at the top, then took the top 81 hits to find 14 NE Tn mutants and assayed these NE mutants alongside JE2 (table 5.2). Some hits were multiple SNPs in the same gene, while others were intergenic or did not have a corresponding NE mutant (i.e. in an essential gene), which reduced the 81 hits to 14 NE mutants. By comparing the lipase activity of JE2 and an NE mutant, it is possible to see if knocking-out the gene does affect lipase activity (table 5.2). We found eight such genes, however two were not investigated further as they were a lipase gene and *agrC*, which is a known regulator of lipase. This left us with six genes to investigate further.

NE mutant	gene (TW20)	gene (USA300)	p-value (Kurskal-Wallis)
NE873	*agrC*	*agrC*	0.003948
NE338	lipase	lipase	0.02497
NE1786	SATW20_19910 (phage antirepressor protein)	SAUSA300_1966	0.01041
NE939	*murA2*	*murA* (2,239,704)	0.01631
NE1868	*mecA*	*mecA*	0.01631
NE181	SATW20_19530	SAUSA300_1936	0.02497
NE1889	*atpH*	*atpH*	0.03737
NE391	*clfB*	*clfB*	0.04495
NE1298	SATW20_19870	SAUSA300_1964	0.05466
NE1826	SATW20_19730	SAUSA300_1953	0.07817
NE499	SATW20_09130	SAUSA300_0817	0.1093
NE585	SATW20_01390	SAUSA300_0130	0.1093
NE1	*ebh*	SAUSA300_1327	0.1495
NE327	*int*	*int*	0.5466

Table 5.2: table showing the 14 NE mutants tested. P-values using Kruskal-Wallis test was calculated for the pNPB (long-chain) substrate, using JE2 for comparison. The Kruskal-Wallis test was used as the distribution of the data showed that a non-parametric test should be used. *agrC* is already a known regulator of lipase, therefore this and the lipase gene will not be investigated further. The bottom 6 NE mutants showed no significant difference in lipase activity compared to JE2, therefore these will also not be investigated further.

Using mutants in other backgrounds to verify the results seen in NE mutants

To verify that the results seen in the NE mutants were specific to the transposon insertion rather than other mutations which may be present in the NE mutants, we used mutants in other genetic backgrounds of *S. aureus*. We had mutants available for *mecA*, as well as a mutant lacking the SCC*mec* locus which carries the *mecA* gene, and *clfB* in other backgrounds, so we characterised them using our lipase assay. BH1CC is a clinical HA-MRSA strain, and LAC is a USA300 CA-MRSA strain[263,264]. The assays revealed no difference between the wild-type and mutants for *mecA* (p=0.09) and *clfB* (p=0.25), so we can conclude that the difference in lipase activity seen in the NE mutants lacking these candidate genes were not due to the deletion of these genes (fig. 5.5).

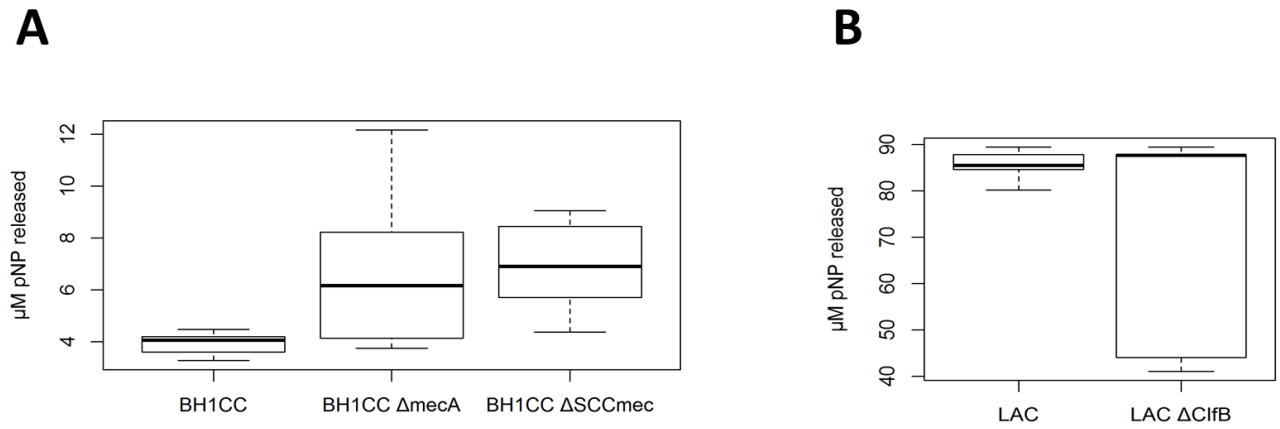


Fig. 5.5: using mutants of *mecA*/SCC*mec* (A) and *clfB* (B) in other backgrounds revealed no difference in lipase activity between the wild-type and mutants (n=6). Therefore, the difference in lipase activity between JE2 and the NE mutants with the above genes knocked-out does not seem to be due to the lack of *clfB* or *mecA*. The box represents the lower and upper quartiles, with the line showing the median value. The whiskers show the lower and upper extremes, and the points outside the whiskers are the outliers.

Complementing the knocked-out genes in the NE mutants

Another way to verify that the reduction in lipase activity seen in the above mutants was due to the knocked-out gene than other mutations is to replace these genes into the NE mutants lacking them by using a tetracycline inducible plasmid vector, pRMC2[243], to see if the re-introduction of the gene of interest restores the lipase activity. First, primers were designed to amplify a fragment more than the whole gene, and had aaa/ttt and a restriction site of choice added to the 5' end. Two restriction sites were chosen to reduce the chances of self-ligation. The vector was isolated from *E. coli* DH5α, and meanwhile, PCR was carried out using JE2 DNA and the above primers to amplify the candidate genes. High-fidelity taq polymerase, such as Phusion, is necessary for this PCR reaction to ensure no mutations occur. The PCR products will be the inserts which will be ligated into the vector.

Once the PCR product was verified as being the correct size, the remaining PCR reaction was purified before being digested with two restriction enzymes, at the same time as the vector in a separate reaction. Once the insert and vector were digested, they were then mixed in

varying ratios and ligated. The re-ligated plasmid was transformed first into *E. coli*, then restriction deficient strain RN4220 before transforming into the relevant NE mutant. The transformants were plated out, and if any colonies grew on the plate they were re-grown in broth and the plasmid was extracted. The plasmid was then double digested again using the same two restriction enzymes to check if the insert is there; two bands will be present if there is, one around 6.5kb (pRMC2) and one where the insert should be. With the SAUSA300_1966 and *atpH*, these ligations worked well and we were able to complement the relevant NE mutants. However, the *murA* gene could not be cloned into pRMC2; we tried using varying vector:insert ratios, and had controls to check transformation efficiency, and using electroporation rather than chemical transformation.

Does replacing SAUSA_1966 and *atpH into their respective mutants restore lipase activity?*

With the two NE mutants that we managed to complement, we again looked at the lipase activity; we compared the complemented NE mutant, the NE mutant and JE2. pRMC2 includes a tetracycline inducible promoter, therefore anhydrotetracycline was added to the culture media to induce the expression of the cloned genes. At higher concentrations of anhydrotetracycline, sometimes the strains did not grow well – these were not analysed, as there would be less lipase present due to the lack of growth.

Unfortunately, when we analysed these strains, there was no difference in lipase activity between the NE mutant and the complemented NE mutant (fig. 5.6). Varying concentrations of anhydrotetracycline was used, ranging from 50ng/ml to 400ng/ml, and several different stocks of anhydrotetracycline was also used to ensure that the lack of induction was not due to a particular batch of the inducer. In the absence of the inducer the SAUSA300_1966 complemented NE1786 shows higher lipase activity compared to NE1786 ($p=0.03$) (fig. 5.6 A). Despite obtaining 6 repeats for each inducer concentration, there was no difference in lipase activity between the NE mutant and the complemented NE mutant. Perhaps anhydrotetracycline affects lipase production in some way, or there may have been an issue with the cloning, such as the gene being cloned in the wrong reading frame.

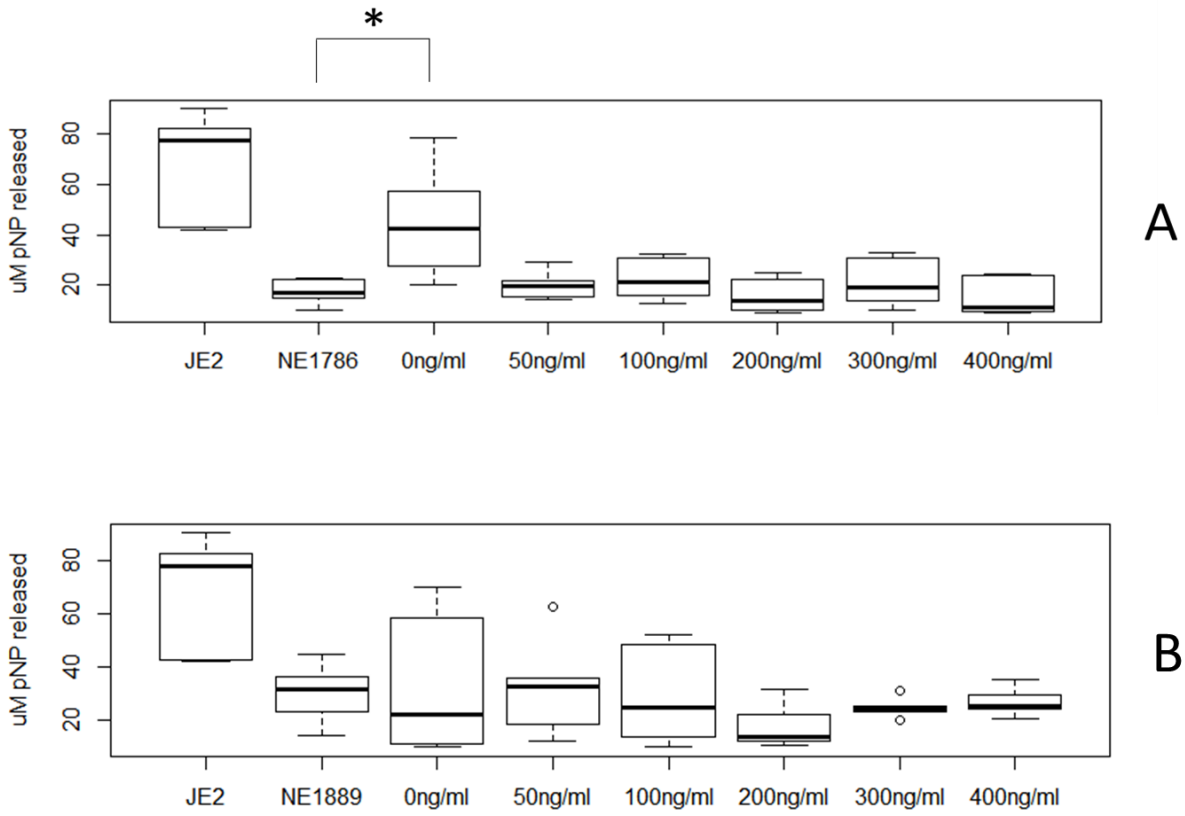


Fig. 5.6: lipase activity of the complemented NE mutants. The complemented NE mutants carry a copy of the knocked-out gene (SAUSA300_1966 (into NE1786, graph A) or *atpH* (into NE1889, graph B)) on a tetracycline inducible plasmid (n=6). The concentrations represent the amount of anhydrotetracycline added to the culture medium. Aside from the SAUSA300_1966 complemented NE1786 in the absence of anhydrotetracycline (0ng/ml, graph A), there is no difference in lipase activity between the NE mutant and the complemented strain (p=0.03, T-test). The box represents the lower and upper quartiles, with the line showing the median value. The whiskers show the lower and upper extremes, and the points outside the whiskers are the outliers.

Discussion

When analysing lipase activity of ST239 MRSA strains, it was noted that it is low compared to the USA300-based strains; however, there was variation between the strains. From this, a GWAS was carried out by Dr Recker and revealed genetic polymorphisms which were statistically associated with a change in lipase production. When we looked through the list of polymorphisms, there were some cases when they occurred in the same gene. Some of the hits were genes in a prophage which had inserted into the lipase precursor gene *geh* or the

phospholipase C precursor gene *hlb*. There is a possibility that in the isolates carrying the polymorphisms within the prophage genes identified by the GWAS, the prophages have inserted into a slightly different location within the *geh* and *hlb* genes. This in turn could give rise to the possibility that a functional fragment of Geh and Hlb could be expressed in these isolates, which would increase lipase activity. Re-sequencing of the ST239 isolates with the current improved technology could help identify if the prophages are in a subtly different position. Alternatively, sequencing which generate long read data (such as Pac Bio or MinION) could be used.

Looking from the most significant, we identified candidate genes and then sought to verify these associations. Of the 14 NE mutants tested, eight were significantly lower in lipase activity compared to the parent strain, JE2. However, two of these were the lipase gene and *agrC*, part of the *accessory gene regulator* quorum sensing system, which is a global regulator of virulence. Therefore, it was unsurprising to find an *agr* gene among the hits, and given that a lipase gene was also among the list of candidate genes shows that the GWAS method used is reliable.

SAUSA300_1966 is an unidentified gene which is listed as "phi77 ORF014-like protein, phage anti-repressor protein" in the GenBank entry. This was shown to be a putative phage-anti-repressor which counteracts the bacteriophage Phi77[248]. This protein may also function as an anti-repressor for lipase, or could be interacting with an unknown repressor of lipase. This would mean that in the presence of the protein, lipase production is either enhanced by the protein itself or enhanced by the protein blocking the action of a lipase repressor.

MurA is part of the cell wall synthesis pathway, and catalyses the first committed step of peptidoglycan biosynthesis; the formation of UDP-*N*-acetylglucosamine-enolpyruvate from UDP-*N*-acetylglucosamine and phosphoenolpyruvate (fig. 5.7)[265]. This protein is also the target of the antibiotic fosfomycin. In the cell wall biosynthesis pathway, a lipid carrier is used to shuttle precursors across the cell membrane, therefore the up-regulation of lipase by *murA* during the stationary phase of growth may be involved in reducing the amount of cell wall biosynthesis by degrading the lipid carrier. The reduction in the number of lipid carriers

available would lead to a reduction in overall cell-wall synthesis, thus slowing down the rate of cell division in the stationary phase. Cell wall lipids could be quantified using methods such as thin-layer chromatography of gas chromatography[255], or by potentially labelling the lipid carriers so that breakdown would either emit or diminish a signal.

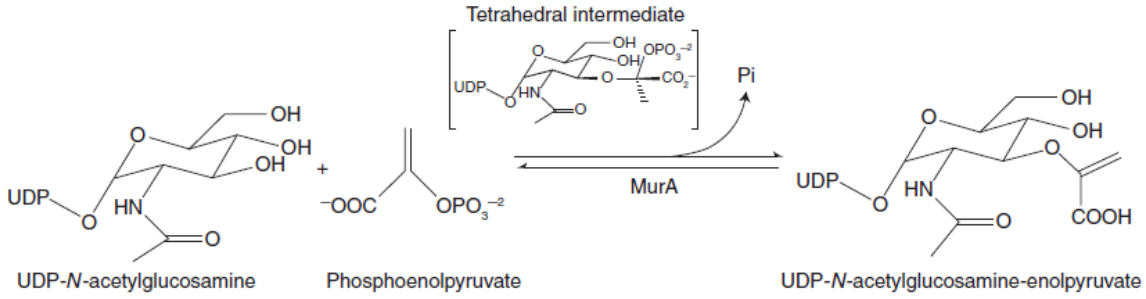


Fig. 5.7: schematic showing the action of MurA, taken from Silver *et al*., Cold Spring Harb Perspect Med 2017[265]. MurA catalysed the formation of UDP-*N*-acetylglucosamine-enolpyruvate from UDP-*N*-acetylglucosamine and phosphoenolpyrivate via a tetrahedral intermediate.

AtpH is the δ subunit of F_1F_0 ATP synthase, which is responsible for providing ATP for the cell[266]. Many chaperones which aid in protein folding has been characterised as ATPases, thus ATP hydrolysis is important for protein folding[267]. Mutations in the δ subunit of ATP synthase has been shown to reduce intracellular ATP concentrations in *E. coli*[266], therefore it is possible that *atpH* is affecting lipase production by interfering with its folding process, due to reduced availability of intracellular ATP. There is a commercial assay kit available from Abcam which quantifies ATP concentrations, so a cell lysate of JE2 and NE1889 could be prepared to be analysed by the above kit.

Attempts to clone the above genes were made, but these were unsuccessful. We managed to successfully ligate the genes into the tetracycline-inducible plasmid pRMC2, but when these were transformed into the corresponding NE mutants there was no difference in lipase activity between the transformed mutants and the NE mutant. A reason could be that the gene was not cloned in the correct reading frame, therefore the primers should be verified, and re-designed if necessary. Interestingly, when *SAUSA300_1966* was replaced into NE1786, in the absence of the inducer (anhydrotetracycline) there was greater lipase activity compared to

NE1786. Perhaps the promoter used in this plasmid is leaky, thus producing some gene product in the absence of the inducer, however if this were the case then it would be expected that the inducer would yield more of the product, thus restoring the lipase activity more. Potentially, anhydrotetracycline may inhibit lipase production, but there is no data to support this at this stage.

In this chapter, we aimed to find novel regulators of lipase production by using a GWAS. It was noted in the GWAS that a lipase gene and a gene from a regulatory system known to affect lipase production, *agrC*, was among the candidate genes from the GWAS. We then sought to verify the other candidate genes by using mutants with these genes knocked-out; this is only possible if the candidate is a non-essential gene. By doing this, we were able to determine if the association identified by the GWAS was a false-positive or an actual association. These revealed six genes, including SAUSA300_1966, *atpH* and *murA*, as genes potentially involved in lipase production.

Future directions

Studies have revealed a potential link between lipase production and virulence in *S. aureus*, as well as link between lipase and host immunity. This would suggest that lipase does have a role in *S. aureus* pathogenicity, and studies like the one in this chapter aims to gain further understanding of how this is regulated and potentially lead to understanting how this affects the overall virulence of *S. aureus*.

Completing the complementation of SAUSA300_1966, *atpH* and *murA*, as well as cloning SAUSA300_1936 is the next step for this research. The *murA* gene was not able to be ligated into pRMC2 despite varying most of the conditions needed for the process. The restriction enzyme site added onto the *murA* forward primer differed from the one added to SAUSA300_1966 and *atpH* primers (EcoRI), as the *murA* gene contains an EcoRI restriction site. The restriction site chosen for the *murA* forward primer was that for SacI and BanII, and as SacI was the available enzyme in the lab this was used for the cloning. Perhaps using BanII in the cloning rather than SacI would improve the outcome of the ligation, so this should be

carried out. If the ligation is still unsuccessful, then the primers should be re-designed, possibly with an alternative restriction enzyme site.

SAUSA300_1966 and *atpH* were successfully ligated into the vector, but their expression was not able to complement the knock-out NE mutants. NE1786 carrying SAUSA300_1966 showed recovery of lipase activity, which was not seen when anhydrotetracycline was added. There is therefore a potential that anhydrotetracycline is interfering with lipase production. To see if this is the case, *agr*+ve strains could be used; these strains could be grown in the presence of anhydrotetracycline to see if this makes a difference to their lipase activity. NE1889 carrying *atpH* did not show a recovery of lipase activity like the NE1786 carrying SAUSA300_1966. The primers could be re-designed to ensure that the complemented strain has the gene in the correct reading frame, and thus is able to recover lipase activity. However this gene may not in fact be associated with lipase activity, so the difference seen between NE1889 and JE2 could be due to other properties of NE1889.

Concluding Remarks

- can we use genome sequences to study bacterial virulence?

Genome Wide Association Studies (GWAS) look for statistical associations between a phenotype and genotypic variations – typically, it is used to identify genetic mutations (Single Nucleotide Polymorphisms (SNPs) for example) which are associated with a disease[239]. An early GWAS in humans compared the genotypes of patients with Age-related Macular Degeneration (AMD) and a control population, and found that a SNP in the complement factor H gene was associated with AMD, leading to the link between AMD and autoimmunity.

The improvements in genome sequencing technology has made it easier to obtain genome sequences for bacteria, and this has led to a rise in the number of genome sequences available[239]. The availability of genome sequences means that it is possible to do GWAS with bacteria, to further investigate how virulence is controlled for example. Some early bacterial GWAS using shotgun sequencing data to be published includes one by Sheppard *et al*., which looked at how *Campylobacter* species adapted to the host, and another by Farhat *et al*. investigated what mutations arose frequently in drug resistant *Mycobacterium tuberculosis*, thereby allowing the identification of genes which promote resistant to anti-tuberculosis drugs[268,269].

The distribution of these genetic polymorphisms is linked to success in GWAS[239]; in humans genetic recombination and chromosomal segregation means that mutations are usually linked to alleles in its vicinity, and the way alleles mix in different genetic backgrounds gives an insight into which of the mutation is causing the phenotypic change as opposed to being passively carried. In bacteria, however, genetic diversity occurs less frequently, by either horizontal gene transfer (movement of mobile genetic elements from one to another), recombination (replacement of DNA) or by recurrent mutation (mutation arising in separate lineages), and these lead to homoplasy; this is where similar genetic loci are found in separate phylogenetic

branches. This is something that needs to be taken into consideration when carrying out bacterial GWAS.

Previous studies in our lab have used GWAS to try and identify novel ways in which virulence is regulated in *Staphylococcus aureus*, particularly Methicillin-Resistant *S. aureus* (MRSA). This includes a study using ST239 MRSA strains to identify new genetic loci which affect toxicity (toxin production) of this lineage, as well as a study which used strains from a single patient and a collection of USA300 strains from various origins[191,246]. The ST239 study and the GWAS carried out using the USA300 strains identified an epistatic association between a SNP in the *ileS* gene and other SNPs at other loci; this means that the presence of the *ileS* SNP does not lead to a change in toxicity, however differences are seen when the presence/absence of the secondary SNPs are taken into account. This particular SNP had been shown to cause resistance to the antibiotic mupirocin, but had not been associated with toxicity before.

The first part of my research focused on uncovering the mechanism behind the association between the SNP in the *ileS* gene causing mupirocin resistance (mupR SNP)[75] and toxin production – between the two MRSA lineages, this was the only SNP in common in this epistasis. This revealed a mechanism by which *S. aureus* balances antibiotic resistance and fitness by modulating virulence, in this case lowering the amount of toxins produced to retain fitness. Studies investigating this particular SNP showed that this was a fitness cost free resistance mechanism[249], and in pure cultures there is no difference seen in the growth rate of isogenic mupirocin sensitive and resistant strains. However, when these strains are forced to compete, fitness is retained only in strains with a functional *agr* system; these *agr* negative strains are not able to produce much toxins. Antibiotic resistance is an increasing problem, and in cases like this mupirocin resistance SNP investigated in my research, the fact that bacteria are able to offset the fitness cost of this resistance mutation means that it is able to be stably maintained in a population; in fact, this mutation is the one most frequently seen in clinical *S. aureus* isolates[75].

The GWAS association seen between mupirocin resistance (mupR) and toxicity was shown to be an epistatis, as stated before. This means that other secondary SNPs are involved in

changing the toxicity of mupirocin resistant strains. This study did not go into details into these secondary SNPs and the role they play in this association; this is a question which still remains. Creating secondary SNP-mutants in isogenic mupirocin sensitive (mupS) and mupR strains will reveal which of these combinations actually lead to a change in toxicity. Perhaps this will help us to further understand how mupR is affecting toxicity. Interestingly, there was no overlap in these secondary SNPs, meaning that the list of secondary SNPs in USA300 and ST239 lineages are different. Does this mean that the GWAS association is lineage specific? – this can be addressed by analysing another sequenced strain collection.

The second part of my research utilised GWAS to find novel genes affecting lipase activity in *S. aureus*. What was first required for this was to create a high throughput approach to analyse lipase activity, which was achieved by adapting an existing assay[261] to be run using a 96 well plate in a plate reader. To validate the assay, an *agr*+ve strain and its isogenic *agr*-ve strain were used, as the *agr* system is known to regulate lipase activity[225]. GWAS revealed genes which were associated with a variation in lipase activity, and these were verified using the Nebraska Transposon (Tn) mutants[170]. To ensure that the results seen was not due to mutations elsewhere, we either used mutants in a different genetic background (where available) or attempted to complement the Tn mutant by introducing the knocked-out gene on an inducible plasmid.

Although we were not able to see difference in lipase activity seen in mutants of *clfB* and *mecA* in other backgrounds, and were not successful in cloning the genes of interest, the GWAS approach was successful in identifying novel genes which seem to have an effect on lipase activity; further research is needed to fully characterise the function of SAUSA300_1966, *murA*, SAUSA300_1936 and *atpH* in modulating lipase activity. For example, when no inducer was added to the SAUSA300_1966 knock-out mutant carrying that gene on pRMC2, there was a significant increase in lipase activity compared to the knock-out mutant without the plasmid; is the inducer, anhydrotetracycline, affecting lipase activity?

These studies demonstrate the potential for GWAS to identify new ways in which virulence of the opportunistic pathogen *S. aureus* is regulated. It is necessary to follow up any significant

hits from the GWAS with experimental work to back-up the finding, as we attempted to do here. Due to the rising problem of antibiotic resistance, it is vital that we understand more about the bacterial pathogens we are facing, and GWAS can be used as a tool by which functions of genes in virulence can be quickly identified. This would allow us to provide new targets for antibiotics which would target the production of virulence factors, and this could be expanded to other pathogens. GWAS may also aid the understanding of virulence in bacteria which have not been well studied, as well as perhaps aiding annotation of genomes through the understanding of the gene's function. This would then increase our understanding of bacteria and facilitate the annotation of the genomes of other organisms. The reduction in cost and time, and the improvements in genome sequencing has seen an increase in the number of sequences available, and I believe there is much we can learn from this information – as demonstrated in this study.

Appendix A

Table 1

Table showing the genes in ST239 (TW20) which translate into proteins with a significantly altered isoleucine (Ile) content, according to the Mote Carlo simulation.

1-31 32-62

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_01700	*cap8K*	Yes	High	SATW20_15510	*mreB*	Yes	High
SATW20_02780	*lytM*	Yes	Low	SATW20_17630	*arsB2*	Yes	High
SATW20_05240		Yes	High	SATW20_20070		Yes	High
SATW20_05720		Yes	High	SATW20_02080		Yes	High
SATW20_06310	*sdrC*	Yes	Low	SATW20_26150		Yes	High
SATW20_06320	*sdrD*	Yes	Low	SATW20_10750		Yes	High
SATW20_06330	*sdrE*	Yes	Low	SATW20_08750		Yes	High
SATW20_06910		Yes	High	SATW20_14080		Yes	High
SATW20_07840		Yes	High	SATW20_16570		Yes	High
SATW20_08220	*tarO*	Yes	High	SATW20_24490		Yes	High
SATW20_08620	*clfA*	Yes	Low	SATW20_05370		Yes	High
SATW20_10490	*atl*	Yes	Low	SATW20_17310		Yes	Low
SATW20_11230	*isdB*	Yes	Low	SATW20_25000		Yes	Low
SATW20_13440		Yes	Low	SATW20_22340	*sceD*	Yes	Low
SATW20_14350	*ebh*	Yes	Low	SATW20_00810		Yes	High
SATW20_17210	*harA*	Yes	Low	SATW20_01480	*tet38*	Yes	High
SATW20_17280		Yes	Low	SATW20_09430		Yes	High
SATW20_20210	*agrC*	Yes	High	SATW20_01430		No	High
SATW20_21710		Yes	Low	SATW20_24720		No	High
SATW20_22950	*fmtB*	Yes	Low	SATW20_13510	*citB*	No	Low
SATW20_23100		Yes	High	SATW20_14780	*ebpS*	No	Low
SATW20_24330		Yes	High	SATW20_01960		No	High
SATW20_27680	*clfB*	Yes	Low	SATW20_28080	*lip*	No	Low
SATW20_27920	*sraP*	Yes	Low	SATW20_20190	*agrB*	No	High
SATW20_00110		Yes	High	SATW20_14370		No	High
SATW20_09780		Yes	High	SATW20_08630		No	Low
SATW20_20150		Yes	Low	SATW20_23400	*hysA*	No	Low
SATW20_26230	*fnbA*	Yes	Low	SATW20_01260	*sirB*	No	High
SATW20_26330		Yes	High	SATW20_09980	*pepB*	No	Low
SATW20_02960		Yes	High	SATW20_06870		No	High
SATW20_08450		Yes	Low	SATW20_16620		No	High

63-103 104-144

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_15830		No	High	SATW20_14360		No	High
SATW20_11240	*isdA*	No	Low	SATW20_16810		No	High
SATW20_05380		No	High	SATW20_15370		No	High
SATW20_22230		No	High	SATW20_26380		No	High
SATW20_10090		No	High	SATW20_11530		No	High
SATW20_08710		No	High	SATW20_27750	*aur*	No	Low
SATW20_01840		No	High	SATW20_18530		No	High
SATW20_23040		No	High	SATW20_06700		No	High
SATW20_02310	*coa*	No	Low	SATW20_04420		No	High
SATW20_26210	*fnbB*	No	Low	SATW20_07230	*fhuB*	No	High
SATW20_05960		No	High	SATW20_27650	*phoB*	No	Low
SATW20_14490	*pbp2*	No	Low	SATW20_15050	*malA*	No	Low
SATW20_07240	*fhuD*	No	High	SATW20_23110		No	High
SATW20_22130		No	High	SATW20_09930	*appC*	No	High
SATW20_01090		No	High	SATW20_28310		No	High
SATW20_01230	*spa*	No	Low	SATW20_26480		No	High
SATW20_24590		No	High	SATW20_25950	*opp-1A*	No	Low
SATW20_20140		No	High	SATW20_04340	*ssb*	No	Low
SATW20_24860	*tcaB*	No	High	SATW20_05220		No	Low
SATW20_25270	*narG*	No	Low	SATW20_25580		No	High
SATW20_07070	*mntB*	No	High	SATW20_19560		No	Low
SATW20_00130		No	High	SATW20_08080	*sstA*	No	High
SATW20_10070		No	High	SATW20_23050		No	High
SATW20_04510		No	Low	SATW20_28050	*icaD*	No	High
SATW20_05201	PSM-α4	No	High	SATW20_07940		No	Low
SATW20_13750	*femB*	No	Low	SATW20_07490		No	High
SATW20_08300		No	Low	SATW20_00660	*tet*	No	High
SATW20_11490		No	High	SATW20_06980		No	High
SATW20_10280		No	High	SATW20_10200		No	High
SATW20_02760		No	High	SATW20_04180		No	High
SATW20_04290		No	High	SATW20_24290		No	High
SATW20_25890		No	High	SATW20_10950		No	High
SATW20_27600		No	High	SATW20_22480	*atpI*	No	High
SATW20_09920	*appB*	No	High	SATW20_01450	*sasD*	No	Low
SATW20_13590		No	High	SATW20_01790		No	High
SATW20_01880		No	High	SATW20_11360		No	High
SATW20_11900	*lspA*	No	High	SATW20_18130		No	High
SATW20_06210		No	Low	SATW20_15990		No	High
SATW20_01680	*cap8I*	No	High	SATW20_28180		No	Low
SATW20_26320		No	High	SATW20_24310		No	High
SATW20_00230	*sasH*	No	Low	SATW20_10710	*ykoC*	No	High

145-185 186-226

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_11840		No	High	SATW20_07800	*saeS*	No	High
SATW20_04700		No	Low	SATW20_18140		No	High
SATW20_17440		No	High	SATW20_14680	*menH*	No	Low
SATW20_00610	*merB*	No	Low	SATW20_28220		No	High
SATW20_14150	*arlS*	No	High	SATW20_24140		No	High
SATW20_16240	*aspS*	No	Low	SATW20_26010		No	High
SATW20_26910		No	High	SATW20_00820		No	High
SATW20_24450		No	High	SATW20_13800		No	High
SATW20_05310		No	High	SATW20_19240		No	High
SATW20_26900	*isaA*	No	Low	SATW20_26880	*ssaA*	No	Low
SATW20_13250		No	High	SATW20_22180	*kdpD*	No	High
SATW20_01510		No	High	SATW20_24820		No	High
SATW20_06120	*rpoB*	No	Low	SATW20_25410	*fmhA*	No	Low
SATW20_27500	*cudB*	No	Low	SATW20_08470	*gap1*	No	Low
SATW20_13210		No	High	SATW20_26360		No	High
SATW20_13760		No	High	SATW20_25790		No	High
SATW20_13870		No	High	SATW20_15350		No	High
SATW20_02530	*ispD1*	No	High	SATW20_11000		No	High
SATW20_11750	*pbpA*	No	Low	SATW20_25430		No	High
SATW20_09880		No	Low	SATW20_01420		No	High
SATW20_23150		No	High	SATW20_02880	*essB*	No	Low
SATW20_13360	*katA*	No	Low	SATW20_23860		No	High
SATW20_15750	*dnaJ*	No	Low	SATW20_09230		No	High
SATW20_01520		No	High	SATW20_13550	*grlB*	No	Low
SATW20_17720		No	High	SATW20_06670		No	High
SATW20_13810		No	High	SATW20_11760	*mraY*	No	High
SATW20_27790		No	High	SATW20_21760		No	Low
SATW20_20090		No	High	SATW20_21770		No	High
SATW20_05900	*nupC*	No	High	SATW20_12070		No	Low
SATW20_19350		No	High	SATW20_26430		No	Low
SATW20_04000		No	High	SATW20_08870		No	High
SATW20_27350		No	Low	SATW20_25820		No	High
SATW20_04080		No	High	SATW20_05400	*gltA*	No	Low
SATW20_01250	*sirC*	No	High	SATW20_22160	*kdpB*	No	High
SATW20_06760		No	High	SATW20_24400		No	Low
SATW20_08110	*sstD*	No	Low	SATW20_27820		No	Low
SATW20_00631		No	High	SATW20_24300		No	Low
SATW20_08360	*lgt*	No	High	SATW20_18600		No	High
SATW20_19110		No	High	SATW20_03180		No	High
SATW20_26520		No	High	SATW20_11930	*pyrP*	No	High
SATW20_03880	*lytA*	No	Low	SATW20_18640		No	High

227-267 268-308

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_26870		No	High	SATW20_03030		No	High
SATW20_23370		No	High	SATW20_01550		No	Low
SATW20_14330		No	High	SATW20_16700	*rpmI*	No	Low
SATW20_06940		No	High	SATW20_20180	*hld*	No	High
SATW20_09760		No	High	SATW20_17880		No	Low
SATW20_15700		No	High	SATW20_02800		No	High
SATW20_07030		No	High	SATW20_21850		No	Low
SATW20_27710	*arcD*	No	High	SATW20_12880		No	High
SATW20_25480	*sbi*	No	Low	SATW20_25520		No	High
SATW20_14000	*msa*	No	High	SATW20_16560	*tag*	No	Low
SATW20_10470		No	High	SATW20_09490	*mnhD*	No	High
SATW20_24480		No	High	SATW20_12690		No	High
SATW20_25510	*hlgB*	No	Low	SATW20_02210		No	High
SATW20_11920	*pyrR*	No	High	SATW20_19260		No	High
SATW20_07260		No	Low	SATW20_23950	*fmhB*	No	Low
SATW20_09590	*glpQ*	No	Low	SATW20_02260		No	Low
SATW20_27910		No	High	SATW20_16530		No	High
SATW20_11870	*ileS*	No	Low	SATW20_08090	*sstB*	No	High
SATW20_17480	*leuS*	No	Low	SATW20_19320	*hlb*	No	Low
SATW20_01920		No	High	SATW20_28390		No	High
SATW20_24110	*modA*	No	Low	SATW20_23980		No	High
SATW20_08070	*rir2*	No	Low	SATW20_04200		No	High
SATW20_25810		No	High	SATW20_10990		No	High
SATW20_07020		No	High	SATW20_26240	*gntP*	No	High
SATW20_08020		No	High	SATW20_07580	*bacA*	No	High
SATW20_10190	*htrA2*	No	Low	SATW20_01220	*lldP1*	No	High
SATW20_21700		No	Low	SATW20_07370	*vraG*	No	High
SATW20_02090		No	High	SATW20_09130		No	High
SATW20_23470		No	High	SATW20_02570	*ispD2*	No	High
SATW20_04940		No	High	SATW20_14850		No	Low
SATW20_15640	*cdd*	No	Low	SATW20_13540		No	High
SATW20_26060		No	Low	SATW20_13830		No	Low
SATW20_25360		No	Low	SATW20_13710	*trpF*	No	High
SATW20_14880		No	Low	SATW20_23300	*lacD*	No	Low
SATW20_28350		No	High	SATW20_13020	*glnA*	No	Low
SATW20_07100		No	High	SATW20_17700		No	High
SATW20_27290		No	Low	SATW20_09530		No	Low
SATW20_25080		No	High	SATW20_10260		No	High
SATW20_23540		No	High	SATW20_27560		No	High
SATW20_03240		No	High	SATW20_23090		No	High
SATW20_23420		No	Low	SATW20_24500		No	High

309-349 350-390

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_23880		No	Low	SATW20_07180	*abcA*	No	High
SATW20_03970		No	High	SATW20_03040		No	High
SATW20_00710		No	High	SATW20_11080		No	High
SATW20_06740		No	High	SATW20_16820	*phoR*	No	High
SATW20_20010		No	Low	SATW20_11860		No	Low
SATW20_14600		No	High	SATW20_13930	*asd*	No	Low
SATW20_19810		No	Low	SATW20_15340		No	High
SATW20_13230		No	High	SATW20_24920		No	High
SATW20_03400		No	Low	SATW20_06520		No	High
SATW20_07410		No	High	SATW20_10970		No	Low
SATW20_10450	*sspA*	No	Low	SATW20_09470	*mnhF*	No	High
SATW20_07770		No	High	SATW20_16160		No	Low
SATW20_27180		No	Low	SATW20_07010		No	High
SATW20_05203		No	High	SATW20_06470		No	High
SATW20_26140		No	High	SATW20_15650		No	High
SATW20_15070		No	Low	SATW20_18210	*hlgB*	No	Low
SATW20_05204		No	High	SATW20_13290		No	High
SATW20_23910		No	High	SATW20_07210		No	High
SATW20_26550		No	High	SATW20_03090		No	High
SATW20_04500		No	High	SATW20_04020	*mepA*	No	High
SATW20_19410	*sea*	No	Low	SATW20_13470		No	High
SATW20_11570		No	High	SATW20_23200	*asp23*	No	Low
SATW20_11610	*otc*	No	Low	SATW20_07390		No	High
SATW20_21730		No	Low	SATW20_16960		No	High
SATW20_06790		No	Low	SATW20_10960		No	High
SATW20_25680		No	High	SATW20_18400		No	High
SATW20_13400		No	Low	SATW20_12910	*glpP*	No	High
SATW20_10130		No	High	SATW20_07290		No	High
SATW20_14820		No	High	SATW20_24980	*lldP2*	No	High
SATW20_20260		No	High	SATW20_16860	*cycA*	No	High
SATW20_23830	*rplD*	No	Low	SATW20_01040		No	High
SATW20_24180		No	High	SATW20_24960		No	High
SATW20_14140	*odhA*	No	Low	SATW20_13730	*trpA*	No	High
SATW20_09730		No	High	SATW20_09750		No	High
SATW20_04400		No	Low	SATW20_07590		No	High
SATW20_19990	*hlb*	No	Low	SATW20_25260	*narH*	No	Low
SATW20_17360		No	High	SATW20_03110		No	High
SATW20_20280		No	High	SATW20_20020		No	Low
SATW20_28290	*cna*	No	Low	SATW20_15490	*sodA*	No	Low
SATW20_09630		No	High	SATW20_02020		No	High
SATW20_15280		No	High	SATW20_02150		No	Low

391-431 432-472

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_01440	*sodM*	No	Low	SATW20_16380	*obg*	No	Low
SATW20_18520		No	High	SATW20_15260		No	High
SATW20_01190		No	High	SATW20_15850		No	Low
SATW20_17740		No	High	SATW20_04550	*pbuX*	No	High
SATW20_14960	*fur*	No	Low	SATW20_07670		No	High
SATW20_16730	*thrS*	No	Low	SATW20_04720		No	Low
SATW20_28370	*vraE*	No	High	SATW20_18950	*gatA*	No	Low
SATW20_28450		No	High	SATW20_27440		No	High
SATW20_21570		No	Low	SATW20_27780		No	High
SATW20_15060	*malR*	No	High	SATW20_01360	*sbnI*	No	Low
SATW20_01000		No	High	SATW20_24630	*hutU*	No	Low
SATW20_24580		No	High	SATW20_25750	*opuCC*	No	Low
SATW20_09250	*lipA*	No	Low	SATW20_22840	*czrB*	No	High
SATW20_21940	*ilvB*	No	Low	SATW20_10870		No	Low
SATW20_24970	*mqo1*	No	Low	SATW20_13880		No	Low
SATW20_07520		No	Low	SATW20_00670		No	Low
SATW20_01170		No	Low	SATW20_12460		No	Low
SATW20_09520	*mnhA*	No	High	SATW20_26650		No	Low
SATW20_13740	*femA*	No	Low	SATW20_04860		No	High
SATW20_00190	*yycG*	No	High	SATW20_01210		No	High
SATW20_21450		No	Low	SATW20_25740	*opuCD*	No	High
SATW20_11120		No	High	SATW20_03010		No	High
SATW20_05910	*ctsR*	No	High	SATW20_07620		No	High
SATW20_08660	*nuc*	No	Low	SATW20_18710		No	Low
SATW20_13580		No	High	SATW20_12210	*recG*	No	Low
SATW20_01770		No	High	SATW20_03770		No	Low
SATW20_23510	*rpsI*	No	Low	SATW20_18900		No	High
SATW20_18940	*gatB*	No	Low	SATW20_16420		No	High
SATW20_04440		No	High	SATW20_03810		No	Low
SATW20_07610	*mgrA*	No	Low	SATW20_13840		No	High
SATW20_17050		No	High	SATW20_16750	*dnaB*	No	Low
SATW20_21461		No	High	SATW20_18460		No	High
SATW20_20460		No	High	SATW20_10360		No	High
SATW20_17890		No	Low	SATW20_04490		No	Low
SATW20_11130		No	Low	SATW20_07650		No	High
SATW20_01270	*sirA*	No	Low	SATW20_23840	*rplC*	No	Low
SATW20_18620		No	Low	SATW20_22960		No	Low
SATW20_22850		No	High	SATW20_07000		No	High
SATW20_01300	*sbnC*	No	Low	SATW20_11320	*pheS*	No	Low
SATW20_00200	*yycH*	No	Low	SATW20_27770		No	Low
SATW20_24760		No	High	SATW20_09480	*mnhE*	No	High

473-513 514-554

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_28330		No	High	SATW20_28410		No	High
SATW20_08200		No	High	SATW20_09620		No	Low
SATW20_08500	*pgm*	No	Low	SATW20_21400		No	High
SATW20_18930		No	High	SATW20_24160	*fhuD*	No	Low
SATW20_05680	*prs*	No	High	SATW20_02330	*fadB*	No	Low
SATW20_05180		No	High	SATW20_00090	*serS*	No	Low
SATW20_21670		No	Low	SATW20_09200		No	High
SATW20_06820	*adhA*	No	Low	SATW20_09460	*mnhG*	No	High
SATW20_01130		No	High	SATW20_10830		No	High
SATW20_21560		No	Low	SATW20_12310	*ffh*	No	Low
SATW20_06170	*fus*	No	Low	SATW20_15450		No	High
SATW20_06550		No	High	SATW20_09020		No	Low
SATW20_11280	*isdF*	No	High	SATW20_25500	*hlgC*	No	Low
SATW20_24890		No	High	SATW20_25470		No	High
SATW20_26890		No	High	SATW20_21340		No	High
SATW20_15480	*pbpF*	No	Low	SATW20_08040		No	High
SATW20_00770		No	High	SATW20_06060	*nusG*	No	Low
SATW20_05110		No	Low	SATW20_25900		No	High
SATW20_00870		No	High	SATW20_25600		No	High
SATW20_01180	*norC*	No	High	SATW20_27550	*nrdD*	No	Low
SATW20_14860		No	Low	SATW20_09120		No	Low
SATW20_28520	*rpmH*	No	Low	SATW20_07700	*norA*	No	High
SATW20_08520		No	High	SATW20_00100		No	High
SATW20_01370		No	High	SATW20_26260	*gntR*	No	Low
SATW20_27810		No	Low	SATW20_10150	*murE*	No	Low
SATW20_25630		No	High	SATW20_00600	*merA*	No	Low
SATW20_21600		No	Low	SATW20_25180	*narT*	No	High
SATW20_00860	*tnpB*	No	Low	SATW20_23230	*opuD2*	No	High
SATW20_01540		No	Low	SATW20_07420		No	High
SATW20_27620		No	High	SATW20_03190		No	High
SATW20_13350		No	High	SATW20_13910		No	Low
SATW20_17230	*acsA*	No	Low	SATW20_24950		No	High
SATW20_20500		No	Low	SATW20_12280		No	Low
SATW20_26190	*sarU*	No	High	SATW20_20380		No	Low
SATW20_14530	*asnS*	No	Low	SATW20_19340		No	High
SATW20_23870		No	High	SATW20_20360		No	Low
SATW20_06080	*rplA*	No	Low	SATW20_09690		No	High
SATW20_12290		No	Low	SATW20_15220		No	High
SATW20_19510		No	High	SATW20_10100		No	High
SATW20_02290		No	High	SATW20_14180		No	High
SATW20_12930	*glpK*	No	Low	SATW20_01070		No	High

555-595 596-636

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_18800		No	High	SATW20_17410		No	Low
SATW20_25940	*opp-1B*	No	High	SATW20_05450		No	Low
SATW20_17120		No	High	SATW20_26530		No	High
SATW20_07990		No	Low	SATW20_12730		No	Low
SATW20_12180	*rpmB*	No	Low	SATW20_00950		No	Low
SATW20_10840		No	High	SATW20_06690		No	High
SATW20_22770		No	Low	SATW20_16850	*citZ*	No	Low
SATW20_11100	*ctaA*	No	High	SATW20_21180		No	Low
SATW20_27740		No	High	SATW20_05150		No	Low
SATW20_02500		No	Low	SATW20_23070		No	High
SATW20_15560		No	High	SATW20_14090		No	Low
SATW20_25590		No	High	SATW20_17500		No	Low
SATW20_22250		No	High	SATW20_12980		No	Low
SATW20_00761		No	High	SATW20_18120		No	Low
SATW20_07830	*saeP*	No	Low	SATW20_15030	*zwf*	No	Low
SATW20_21590		No	Low	SATW20_02600		No	Low
SATW20_07400		No	Low	SATW20_20110		No	High
SATW20_00120		No	Low	SATW20_17460	*sasC*	No	Low
SATW20_23650	*rplO*	No	Low	SATW20_16390	*rpmA*	No	Low
SATW20_09190		No	High	SATW20_02720		No	High
SATW20_15180		No	Low	SATW20_24690		No	Low
SATW20_25690		No	High	SATW20_22980		No	High
SATW20_24600		No	Low	SATW20_07200		No	High
SATW20_03630		No	Low	SATW20_25930	*opp-1C*	No	High
SATW20_14290		No	Low	SATW20_05020		No	Low
SATW20_08010		No	High	SATW20_14260		No	High
SATW20_21290		No	Low	SATW20_20630		No	Low
SATW20_02340	*fadD*	No	Low	SATW20_23020		No	High
SATW20_20120	*groEL*	No	Low	SATW20_01670	*cap8H*	No	High
SATW20_01560		No	Low	SATW20_22510	*glyA*	No	Low
SATW20_18230		No	High	SATW20_12360		No	High
SATW20_00480		No	Low	SATW20_10420	*menB*	No	Low
SATW20_06660		No	High	SATW20_25440		No	Low
SATW20_16030		No	High	SATW20_15710	*rpsU*	No	Low
SATW20_15390		No	High	SATW20_26750		No	High
SATW20_24830		No	High	SATW20_13170		No	High
SATW20_23270	*lacG*	No	Low	SATW20_15190		No	Low
SATW20_19040	*scpA*	No	Low	SATW20_16760		No	Low
SATW20_09510	*mnhB*	No	High	SATW20_07130	*tagG*	No	High
SATW20_08370		No	High	SATW20_22720	*luxS*	No	Low
SATW20_04210		No	High	SATW20_06160	*rpsG*	No	Low

637-676 677-716

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_12440	*topA*	No	Low	SATW20_00400	*tnpB1*	No	Low
SATW20_04850		No	High	SATW20_24540		No	High
SATW20_01140	*plc*	No	Low	SATW20_16510	*tnpB2*	No	Low
SATW20_12940	*glpD*	No	Low	SATW20_19120		No	High
SATW20_15310		No	Low	SATW20_00620		No	Low
SATW20_16360	*ruvA*	No	Low	SATW20_27240		No	Low
SATW20_11220	*rpmF*	No	Low	SATW20_23460		No	High
SATW20_02640	*lrgA*	No	High	SATW20_02890	*essC*	No	High
SATW20_24710		No	High	SATW20_28480	*gidB*	No	Low
SATW20_17420		No	Low	SATW20_18630		No	High
SATW20_11430	*sdhB*	No	Low	SATW20_23530		No	Low
SATW20_16540	*folC*	No	High	SATW20_09960		No	Low
SATW20_13460		No	High	SATW20_03890		No	High
SATW20_09770		No	High	SATW20_16630	*hemA*	No	High
SATW20_18510	*glnQ*	No	Low	SATW20_20700	*aacA-aphD*	No	High
SATW20_12920	*glpF*	No	High	SATW20_13490	*mscL*	No	High
SATW20_13570		No	High	SATW20_07270		No	High
SATW20_18450	*citG*	No	Low	SATW20_10760		No	Low
SATW20_13220		No	High	SATW20_21550		No	Low
SATW20_24460		No	High	SATW20_12470	*hslV*	No	Low
SATW20_03150		No	High	SATW20_02610	*scdA*	No	Low
SATW20_03660		No	Low	SATW20_06750		No	High
SATW20_13100		No	High	SATW20_23670	*rpsE*	No	Low
SATW20_10430	*sspC*	No	High	SATW20_13380		No	Low
SATW20_11070		No	Low	SATW20_14580		No	Low
SATW20_11410	*sdhC*	No	High	SATW20_24340		No	Low
SATW20_10440	*sspB*	No	Low	SATW20_07440		No	Low
SATW20_22090		No	Low	SATW20_05470		No	Low
SATW20_19780		No	Low	SATW20_11420	*sdhA*	No	Low
SATW20_25300	*nasD*	No	Low	SATW20_27110	*blaR1*	No	High
SATW20_03430		No	Low	SATW20_25070		No	High
SATW20_14900	*srrA*	No	Low	SATW20_02490		No	High
SATW20_04130		No	High	SATW20_24670		No	High
SATW20_19210		No	High	SATW20_18730		No	Low
SATW20_24570		No	Low	SATW20_05850	*lysS*	No	Low
SATW20_27040		No	High	SATW20_26440		No	Low
SATW20_05390		No	High	SATW20_02040		No	High
SATW20_04030	*mepB*	No	Low	SATW20_18350		No	Low
SATW20_27450	*fda*	No	Low	SATW20_26680	*mvaS*	No	Low
SATW20_09220		No	Low	SATW20_01630	*capD*	No	High

717-757 758-798

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_02270		No	Low	SATW20_16130	*alaS*	No	Low
SATW20_23900	*glcU*	No	High	SATW20_05120		No	Low
SATW20_19300		No	Low	SATW20_00880		No	High
SATW20_17140		No	High	SATW20_03800		No	Low
SATW20_01870		No	Low	SATW20_03510		No	Low
SATW20_20240	*scrB*	No	Low	SATW20_07910		No	High
SATW20_19290		No	High	SATW20_14510		No	Low
SATW20_17900		No	Low	SATW20_19840		No	High
SATW20_13340		No	Low	SATW20_08530	*secG*	No	High
SATW20_20760	*aphA*	No	Low	SATW20_11630		No	High
SATW20_08790		No	Low	SATW20_24090	*modC*	No	High
SATW20_04810		No	Low	SATW20_03670		No	Low
SATW20_04140		No	High	SATW20_27070		No	High
SATW20_02850	*esaA*	No	Low	SATW20_04970		No	High
SATW20_04430		No	Low	SATW20_19880		No	High
SATW20_07980		No	High	SATW20_19160		No	Low
SATW20_03840		No	Low	SATW20_11010		No	Low
SATW20_20670		No	Low	SATW20_22600		No	Low
SATW20_15380		No	High	SATW20_13480		No	High
SATW20_04300		No	High	SATW20_25140	*gltT*	No	High
SATW20_09400		No	Low	SATW20_18170		No	High
SATW20_12080		No	High	SATW20_00050	*gyrB*	No	Low
SATW20_04660		No	Low	SATW20_23930		No	High
SATW20_17390		No	Low	SATW20_06570	*pta*	No	Low
SATW20_06730		No	High	SATW20_12770	*pgsA*	No	High
SATW20_28260		No	Low	SATW20_28120	*hisH*	No	High
SATW20_18980		No	Low	SATW20_13140		No	High
SATW20_22560	*prfA*	No	Low	SATW20_21810		No	Low
SATW20_14830		No	Low	SATW20_13900	*cvfB*	No	Low
SATW20_19380	*sak*	No	Low	SATW20_06180	*tuf*	No	Low
SATW20_06240		No	Low	SATW20_19820		No	High
SATW20_06490	*thiD*	No	Low	SATW20_03390		No	High
SATW20_26620	*cidA*	No	High	SATW20_08640		No	Low
SATW20_12060	*priA*	No	Low	SATW20_08320		No	High
SATW20_23440		No	Low	SATW20_14840		No	Low
SATW20_01850		No	Low	SATW20_27950		No	Low
SATW20_19500		No	Low	SATW20_20870		No	High
SATW20_22700		No	Low	SATW20_15820		No	High
SATW20_24430		No	High	SATW20_28170	*hisZ*	No	Low
SATW20_14250	*msrA1*	No	Low	SATW20_25670		No	High
SATW20_25250	*narJ*	No	Low	SATW20_18990	*lig*	No	Low

799-839 840-880

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_15110		No	Low	SATW20_15580	*dnaG*	No	Low
SATW20_21580		No	Low	SATW20_23640	*secY*	No	High
SATW20_04220	*thl*	No	Low	SATW20_18340		No	High
SATW20_24100	*modB*	No	High	SATW20_28300		No	High
SATW20_05540		No	Low	SATW20_07380		No	High
SATW20_08100	*sstC*	No	High	SATW20_27340		No	High
SATW20_21100		No	Low	SATW20_17100	*rpsD*	No	Low
SATW20_17090		No	High	SATW20_08060	*rir1*	No	Low
SATW20_02650	*lrgB*	No	High	SATW20_06680		No	High
SATW20_18300		No	High	SATW20_27720	*arcB*	No	Low
SATW20_27930		No	Low	SATW20_27360	*panD*	No	High
SATW20_12850		No	High	SATW20_06620		No	Low
SATW20_10050		No	High	SATW20_15470	*rpmG1*	No	Low
SATW20_03100		No	High	SATW20_21010		No	Low
SATW20_18970	*putP*	No	High	SATW20_15760	*dnaK*	No	Low
SATW20_26690		No	Low	SATW20_14010	*cspA*	No	Low
SATW20_11650		No	High	SATW20_05560		No	Low
SATW20_21620		No	Low	SATW20_04670		No	Low
SATW20_07170	*pbp4*	No	Low	SATW20_02630	*lytR*	No	High
SATW20_23130		No	High	SATW20_01820		No	High
SATW20_20430		No	Low	SATW20_00580		No	High
SATW20_13860		No	High	SATW20_05950	*radA*	No	Low
SATW20_26130		No	High	SATW20_26850		No	Low
SATW20_18470		No	Low	SATW20_21380		No	High
SATW20_23770	*rplP*	No	Low	SATW20_21330		No	Low
SATW20_16040		No	High	SATW20_25390		No	Low
SATW20_00990		No	High	SATW20_18370		No	Low
SATW20_24530		No	High	SATW20_27470		No	Low
SATW20_19620		No	Low	SATW20_15360		No	High
SATW20_20350		No	High	SATW20_03830		No	High
SATW20_18220	*hlgA*	No	Low	SATW20_19790		No	Low
SATW20_28110	*hisA*	No	High	SATW20_16550	*valS*	No	Low
SATW20_09350	*dltD*	No	Low	SATW20_09840	*oppB*	No	High
SATW20_23590	*rpsK*	No	Low	SATW20_06230		No	Low
SATW20_22460	*atpE*	No	High	SATW20_03420		No	Low
SATW20_14450		No	Low	SATW20_17680		No	High
SATW20_01410		No	High	SATW20_08380		No	Low
SATW20_17820		No	Low	SATW20_05330		No	Low
SATW20_01280	*sbnA*	No	High	SATW20_21960	*ilvC*	No	Low
SATW20_02810		No	High	SATW20_05420		No	High
SATW20_07950		No	Low	SATW20_02460		No	High

881-887

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SATW20_15300		No	High
SATW20_16190		No	Low
SATW20_09180		No	Low
SATW20_05800		No	Low
SATW20_15920	*aroE*	No	High
SATW20_12510	*tsf*	No	Low
SATW20_09990		No	High

Table 2

Table showing the genes in USA300 (FPR3757) which translate into proteins with a significantly altered isoleucine (Ile) content, according to the Mote Carlo simulation.

1-33 34-66

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SAUSA300_0161	*cap5J*	Y	High	SAUSA300_1613		Y	High
SAUSA300_0162	*cap5K*	Y	High	SAUSA300_1718	*arsB*	Y	High
SAUSA300_0428		Y	High	SAUSA300_0754		Y	Low
SAUSA300_0546	*sdrC*	Y	Low	SAUSA300_0881		Y	High
SAUSA300_0547	*sdrD*	Y	Low	SAUSA300_0980		Y	High
SAUSA300_0548	*sdrE*	Y	Low	SAUSA300_1979		Y	High
SAUSA300_0731		Y	High	SAUSA300_1985	*sdrH*	Y	Low
SAUSA300_0772	*clfA*	Y	Low	SAUSA300_2266		Y	High
SAUSA300_0955	*atl*	Y	Low	SAUSA300_2434		Y	High
SAUSA300_1028		Y	Low	SAUSA300_2441	*fnbA*	Y	Low
SAUSA300_1239	*tkt*	Y	Low	SAUSA300_0442		Y	High
SAUSA300_1327		Y	Low	SAUSA300_1300	*brnQ (3)*	Y	High
SAUSA300_1393		Y	Low	SAUSA300_1687		Y	Low
SAUSA300_1677		Y	Low	SAUSA300_0224	*coa*	Y	Low
SAUSA300_1684		Y	Low	SAUSA300_2315		Y	Low
SAUSA300_1991	*agrC*	Y	High	SAUSA300_0846		Y	High
SAUSA300_2109	*fmtB*	Y	Low	SAUSA300_2051		Y	Low
SAUSA300_2134		Y	High	SAUSA300_1370	*ebpS*	Y	Low
SAUSA300_2252		Y	High	SAUSA300_0773		N	Low
SAUSA300_2565	*clfB*	Y	Low	SAUSA300_0116	*sirB*	N	High
SAUSA300_2589		Y	Low	SAUSA300_0091		N	High
SAUSA300_0083		Y	High	SAUSA300_0291		N	High
SAUSA300_0134		Y	High	SAUSA300_1246	*acnA*	N	Low
SAUSA300_0270	*lytM*	Y	Low	SAUSA300_2287		N	High
SAUSA300_0276		Y	High	SAUSA300_1329		N	High
SAUSA300_2451		Y	High	SAUSA300_0902	*pepF*	N	Low
SAUSA300_0011		Y	High	SAUSA300_2161	*hysA*	N	Low
SAUSA300_0113		Y	Low	SAUSA300_1989	*agrB*	N	High
SAUSA300_0482		Y	High	SAUSA300_2440	*fnbB*	N	Low
SAUSA300_0603		Y	High	SAUSA300_1547		N	High
SAUSA300_0694		Y	High	SAUSA300_1618	*hemX*	N	High
SAUSA300_0784		Y	High	SAUSA300_2603	*lip*	N	Low
SAUSA300_1515		Y	High	SAUSA300_2128		N	High

67-105 106-144

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SAUSA300_0073		N	Low	SAUSA300_0895	*oppB (2)*	N	High
SAUSA300_0320		N	Low	SAUSA300_0176		N	High
SAUSA300_0599		N	High	SAUSA300_2561	*phoB*	N	Low
SAUSA300_0110		N	High	SAUSA300_0634	*fhuB*	N	High
SAUSA300_0201		N	High	SAUSA300_2624		N	High
SAUSA300_1029		N	Low	SAUSA300_0109		N	High
SAUSA300_2040		N	High	SAUSA300_0374		N	High
SAUSA300_1341	*pbp2*	N	Low	SAUSA300_2411	*opp-1A*	N	Low
SAUSA300_0139		N	High	SAUSA300_1809		N	High
SAUSA300_0268		N	High	SAUSA300_0583		N	High
SAUSA300_2343		N	Low	SAUSA300_2135		N	High
SAUSA300_2110	*fmtB (2)*	N	Low	SAUSA300_1456		N	Low
SAUSA300_0443		N	High	SAUSA300_0426		N	Low
SAUSA300_0913		N	High	SAUSA300_0143	*phnE (2)*	N	High
SAUSA300_2030		N	High	SAUSA300_0367	*ssb*	N	Low
SAUSA300_0188	*brnQ*	N	High	SAUSA300_2628	*rarD*	N	High
SAUSA300_0635	*fhuG*	N	High	SAUSA300_2466		N	High
SAUSA300_0619		N	High	SAUSA300_0896	*oppC (2)*	N	High
SAUSA300_2274		N	High	SAUSA300_1254		N	High
SAUSA300_0013		N	High	SAUSA300_1938		N	Low
SAUSA300_2301	*tcaB*	N	High	SAUSA300_0932		N	High
SAUSA300_0136		N	Low	SAUSA300_2374		N	High
SAUSA300_1984		N	High	SAUSA300_0660		N	High
SAUSA300_0911		N	High	SAUSA300_0610		N	High
SAUSA300_0512		N	High	SAUSA300_0703		N	Low
SAUSA300_1270	*femB*	N	Low	SAUSA300_0718		N	High
SAUSA300_0383		N	Low	SAUSA300_0351		N	High
SAUSA300_1637		N	High	SAUSA300_2129		N	High
SAUSA300_0739		N	Low	SAUSA300_0924		N	High
SAUSA300_2405		N	High	SAUSA300_1041		N	High
SAUSA300_1054		N	High	SAUSA300_0025		N	Low
SAUSA300_0934		N	High	SAUSA300_1360	*ubiE*	N	Low
SAUSA300_2556		N	High	SAUSA300_1501		N	High
SAUSA300_2450		N	High	SAUSA300_0180		N	High
SAUSA300_0536		N	Low	SAUSA300_1760	*epiG*	N	High
SAUSA300_0362		N	High	SAUSA300_1000	*potB*	N	High
SAUSA300_1089	*lspA*	N	High	SAUSA300_2248		N	High
	PSMα-4	N	High	SAUSA300_0074	*opp-3B*	N	High
SAUSA300_2572	*aur*	N	Low	SAUSA300_1561		N	High

145-185 186-226

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SAUSA300_2456		N	High	SAUSA300_0744	*lgt*	N	High
SAUSA300_1217		N	High	SAUSA300_2233		N	High
SAUSA300_1084		N	High	SAUSA300_1534		N	High
SAUSA300_0127		N	High	SAUSA300_1761	*epiE*	N	High
SAUSA300_0977		N	High	SAUSA300_2503		N	Low
SAUSA300_2250	*nhaC*	N	High	SAUSA300_0282		N	Low
SAUSA300_1586	*aspS*	N	Low	SAUSA300_0756	*gap*	N	Low
SAUSA300_2506	*isaA*	N	Low	SAUSA300_2417		N	High
SAUSA300_1700		N	High	SAUSA300_2033	*kdpB*	N	High
SAUSA300_1307	*arlS*	N	High	SAUSA300_1250	*parE*	N	Low
SAUSA300_0527	*ropB*	N	Low	SAUSA300_1910		N	High
SAUSA300_2602	*icaC*	N	High	SAUSA300_2454		N	High
SAUSA300_1075	*pbpA*	N	Low	SAUSA300_2297		N	High
SAUSA300_2545	*betA*	N	Low	SAUSA300_2576		N	High
SAUSA300_2262		N	High	SAUSA300_2206		N	High
SAUSA300_1221		N	High	SAUSA300_0445	*gltB*	N	Low
SAUSA300_0436		N	High	SAUSA300_2358		N	High
SAUSA300_0142	*phnE*	N	High	SAUSA300_1106		N	Low
SAUSA300_0891	*oppA*	N	Low	SAUSA300_2461		N	Low
SAUSA300_1282	*pstC*	N	High	SAUSA300_0171		N	High
SAUSA300_1271		N	High	SAUSA300_2579		N	Low
SAUSA300_1232		N	Low	SAUSA300_2395		N	High
SAUSA300_1328		N	High	SAUSA300_2258		N	Low
SAUSA300_1539	*dnaJ*	N	Low	SAUSA300_0581		N	High
SAUSA300_1276	*opp-2B*	N	High	SAUSA300_0827		N	High
SAUSA300_2531		N	Low	SAUSA300_0797		N	High
SAUSA300_0215		N	High	SAUSA300_2249	*ssaA*	N	Low
SAUSA300_0721		N	Low	SAUSA300_1275		N	High
SAUSA300_0506	*nupC*	N	High	SAUSA300_1005		N	High
SAUSA300_0341		N	High	SAUSA300_1076	*mraY*	N	High
SAUSA300_1500		N	High	SAUSA300_2398		N	High
SAUSA300_2614		N	Low	SAUSA300_2174		N	High
SAUSA300_0588		N	High	SAUSA300_0313		N	High
SAUSA300_0333		N	High	SAUSA300_2364	*sbi*	N	Low
SAUSA300_2470	*sdaAB*	N	High	SAUSA300_0115	*sirC*	N	High
SAUSA300_2139		N	High	SAUSA300_0245		N	High
SAUSA300_0690	*saeS*	N	High	SAUSA300_0712		N	High
SAUSA300_2356	*fmhA*	N	Low	SAUSA300_1847		N	High
SAUSA300_2507		N	High	SAUSA300_1397		N	Low
SAUSA300_1897		N	High	SAUSA300_1092	*pyrP*	N	High
SAUSA300_0297		N	High	SAUSA300_2035	*kdpD*	N	High

227-267 268-308

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SAUSA300_0606		N	High	SAUSA300_1612	*tag*	N	Low
SAUSA300_0064	*arcD*	N	High	SAUSA300_1918		N	Low
SAUSA300_1851		N	High	SAUSA300_2568	*arcD (2)*	N	High
SAUSA300_1324		N	High	SAUSA300_2214		N	Low
SAUSA300_0637		N	Low	SAUSA300_0272		N	High
SAUSA300_0879		N	High	SAUSA300_0330	*ulaA*	N	High
SAUSA300_1087	*ileS*	N	Low	SAUSA300_1187		N	High
SAUSA300_0615		N	High	SAUSA300_1920	*chs*	N	Low
SAUSA300_2367	*hlgB*	N	Low	SAUSA300_2368		N	High
SAUSA300_0862	*glpQ*	N	Low	SAUSA300_1168		N	High
SAUSA300_0220	*pflB*	N	Low	SAUSA300_0687		N	High
SAUSA300_1091	*pyrR*	N	High	SAUSA300_1461		N	High
SAUSA300_2230	*modA*	N	Low	SAUSA300_0852	*mnhD*	N	High
SAUSA300_1704	*leuS*	N	Low	SAUSA300_1004		N	High
SAUSA300_1294		N	High	SAUSA300_1912		N	High
SAUSA300_1378		N	Low	SAUSA300_1609		N	High
SAUSA300_2618		N	High	SAUSA300_2217		N	High
SAUSA300_0953		N	High	SAUSA300_2637		N	High
SAUSA300_0133		N	High	SAUSA300_0353		N	High
SAUSA300_2265		N	High	SAUSA300_1278	*pepF (2)*	N	Low
SAUSA300_2588	*secY (2)*	N	High	SAUSA300_0719		N	High
SAUSA300_0717	*nrdF*	N	Low	SAUSA300_1739		N	Low
SAUSA300_0614		N	High	SAUSA300_0184	*argB*	N	High
SAUSA300_2351		N	Low	SAUSA300_1201	*glnA*	N	Low
SAUSA300_1440		N	Low	SAUSA300_2152	*lacD*	N	Low
SAUSA300_2397		N	High	SAUSA300_2442	*gntP*	N	High
SAUSA300_2525		N	Low	SAUSA300_1266	*trpF*	N	High
SAUSA300_1528	*cdd*	N	Low	SAUSA300_0112	*lctP*	N	High
SAUSA300_0146		N	Low	SAUSA300_0856		N	Low
SAUSA300_2163		N	Low	SAUSA300_0817		N	High
SAUSA300_0393		N	High	SAUSA300_0304		N	High
SAUSA300_0923	*htrA*	N	Low	SAUSA300_1724		N	High
SAUSA300_0202		N	High	SAUSA300_2208	*topB*	N	Low
SAUSA300_2168		N	High	SAUSA300_0249	*ispD*	N	High
SAUSA300_1626	*rpmI*	N	Low	SAUSA300_1249		N	High
SAUSA300_0648		N	High	SAUSA300_1974		N	Low
SAUSA300_0622		N	High	SAUSA300_1436		N	Low
SAUSA300_0050		N	High	SAUSA300_1483		N	Low
SAUSA300_0937		N	High	SAUSA300_0319		N	High
SAUSA300_2323	*cobI*	N	High	SAUSA300_1352		N	High
SAUSA300_1988		N	High	SAUSA300_0951	*sspA*	N	Low

309-349 350-390

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SAUSA300_1410		N	Low	SAUSA300_1013		N	High
SAUSA300_2211		N	High	SAUSA300_0120	*sbnC*	N	Low
SAUSA300_0652		N	High	SAUSA300_1499	*aroK*	N	High
SAUSA300_1219		N	High	SAUSA300_2133		N	High
SAUSA300_1458		N	Low	SAUSA300_0565		N	High
SAUSA300_2632		N	High	SAUSA300_0850	*mnhF*	N	High
SAUSA300_1062	*argF*	N	Low	SAUSA300_2307		N	High
SAUSA300_1236		N	Low	SAUSA300_2479	*cidA*	N	High
SAUSA300_0105		N	Low	SAUSA300_1529	*dgkA*	N	High
SAUSA300_2502		N	High	SAUSA300_0613		N	High
SAUSA300_0591		N	Low	SAUSA300_2142	*asp23*	N	Low
SAUSA300_1306	*sucA*	N	Low	SAUSA300_1253	*glcT*	N	High
SAUSA300_0062	*arcB*	N	Low	SAUSA300_2312	*mqo*	N	Low
SAUSA300_1578	*mnmA*	N	Low	SAUSA300_1225		N	High
SAUSA300_1692		N	High	SAUSA300_0632		N	High
SAUSA300_0382		N	High	SAUSA300_1242	*sbcD*	N	High
SAUSA300_0586		N	High	SAUSA300_0917		N	High
SAUSA300_2203	*rplD*	N	Low	SAUSA300_2207		N	High
SAUSA300_0372		N	Low	SAUSA300_2342	*narH*	N	Low
SAUSA300_2384		N	High	SAUSA300_0335		N	High
SAUSA300_1374		N	High	SAUSA300_1651		N	High
SAUSA300_0669		N	High	SAUSA300_1001	*potC*	N	High
SAUSA300_0344		N	Low	SAUSA300_1796		N	High
SAUSA300_1287	*asd*	N	Low	SAUSA300_0650		N	High
SAUSA300_0866		N	High	SAUSA300_1975		N	Low
SAUSA300_0081		N	High	SAUSA300_0135		N	Low
SAUSA300_1996	*amt*	N	High	SAUSA300_0209		N	Low
SAUSA300_1086		N	Low	SAUSA300_1513		N	Low
	PSMα-1	N	High	SAUSA300_1629	*thrS*	N	Low
SAUSA300_0876		N	High	SAUSA300_1642		N	High
SAUSA300_0942		N	High	SAUSA300_2313		N	High
SAUSA300_0290		N	High	SAUSA300_2311		N	High
SAUSA300_2552		N	High	SAUSA300_0670		N	High
SAUSA300_1401		N	Low	SAUSA300_1268	*trpA*	N	High
	PSMα-2	N	High	SAUSA300_1638	*phoR*	N	High
SAUSA300_1002	*potD*	N	Low	SAUSA300_2007	*ilvB*	N	Low
SAUSA300_1493		N	High	SAUSA300_1448		N	Low
SAUSA300_1998		N	High	SAUSA300_2171	*rpsI*	N	Low
SAUSA300_2436		N	Low	SAUSA300_0107		N	High
SAUSA300_0630		N	High	SAUSA300_0878		N	High
SAUSA300_1768	*lukD*	N	Low	SAUSA300_0829	*lipA*	N	Low

391-431 432-472

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SAUSA300_0194		N	High	SAUSA300_0035		N	Low
SAUSA300_0306	*brnQ (2)*	N	High	SAUSA300_1283	*pstS*	N	Low
SAUSA300_0663		N	Low	SAUSA300_0531		N	Low
SAUSA300_1726		N	High	SAUSA300_2482		N	Low
SAUSA300_1808		N	High	SAUSA300_1120	*recG*	N	Low
SAUSA300_0776	*nuc*	N	Low	SAUSA300_1858		N	Low
SAUSA300_2642		N	High	SAUSA300_2575		N	High
SAUSA300_0568		N	High	SAUSA300_2099		N	High
SAUSA300_1549		N	Low	SAUSA300_0640		N	High
SAUSA300_1880	*gatB*	N	Low	SAUSA300_1631		N	Low
SAUSA300_1457	*malR*	N	High	SAUSA300_0381		N	Low
SAUSA300_2472		N	High	SAUSA300_2390	*opuCd*	N	High
SAUSA300_0672		N	Low	SAUSA300_0111		N	High
SAUSA300_0507	*ctsR*	N	High	SAUSA300_1015	*ctaA*	N	High
SAUSA300_1017		N	High	SAUSA300_2204	*rplC*	N	Low
SAUSA300_1190	*glpP*	N	High	SAUSA300_0673		N	High
SAUSA300_1740		N	Low	SAUSA300_2111	*glmM*	N	Low
SAUSA300_0128		N	High	SAUSA300_1037	*pheS*	N	Low
SAUSA300_0855	*mnhA*	N	High	SAUSA300_1877		N	High
SAUSA300_1849	*mutY*	N	Low	SAUSA300_1279	*phoU*	N	High
SAUSA300_0021		N	High	SAUSA300_1802		N	High
SAUSA300_1018		N	Low	SAUSA300_1604	*mreD*	N	High
SAUSA300_0376		N	High	SAUSA300_0759	*gpmI*	N	Low
SAUSA300_0117	*sirA*	N	Low	SAUSA300_0222		N	High
SAUSA300_1600	*obgE*	N	Low	SAUSA300_0532	*fusA*	N	Low
SAUSA300_2273		N	High	SAUSA300_0676		N	High
SAUSA300_1881	*gatA*	N	Low	SAUSA300_0594	*adh*	N	Low
SAUSA300_1660		N	High	SAUSA300_0612		N	High
SAUSA300_2278	*hutU*	N	Low	SAUSA300_0094		N	High
SAUSA300_1663		N	High	SAUSA300_0851	*mnhE*	N	High
SAUSA300_2578		N	Low	SAUSA300_0408		N	Low
SAUSA300_2391	*opuCc*	N	Low	SAUSA300_0868	*spsB*	N	Low
SAUSA300_0126		N	Low	SAUSA300_0729		N	High
SAUSA300_2100		N	High	SAUSA300_0478	*prs*	N	High
SAUSA300_2291	*gltS*	N	High	SAUSA300_1033		N	High
SAUSA300_1491		N	High	SAUSA300_0098		N	High
SAUSA300_1145	*xerC*	N	Low	SAUSA300_1379		N	Low
SAUSA300_0992		N	Low	SAUSA300_2648	*rpmH*	N	Low
SAUSA300_0022		N	Low	SAUSA300_1679	*acsA*	N	Low
SAUSA300_0387	*pbuX*	N	High	SAUSA300_0145		N	Low
SAUSA300_2237		N	High	SAUSA300_2304		N	High

473-513 514-554

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SAUSA300_1128	*ftsY*	N	Low	SAUSA300_0708	*hisC*	N	Low
SAUSA300_1345	*asnC*	N	Low	SAUSA300_2333	*narK*	N	High
SAUSA300_0523	*rplA*	N	Low	SAUSA300_2540		N	Low
SAUSA300_0761		N	High	SAUSA300_0314		N	High
SAUSA300_1192	*glpK*	N	Low	SAUSA300_0872		N	High
SAUSA300_0263	*rbsD*	N	High	SAUSA300_1474		N	High
SAUSA300_0865	*pgi*	N	Low	SAUSA300_1117	*rpmB*	N	Low
SAUSA300_0147		N	Low	SAUSA300_1310		N	High
SAUSA300_2379		N	High	SAUSA300_0242	*gutB*	N	Low
SAUSA300_2235		N	Low	SAUSA300_2092	*dps*	N	Low
SAUSA300_0009	*serS*	N	Low	SAUSA300_0914		N	High
SAUSA300_2558	*nsaS*	N	High	SAUSA300_0066	*argR*	N	High
SAUSA300_1130	*ffh*	N	Low	SAUSA300_2410		N	High
SAUSA300_2438	*sarU*	N	High	SAUSA300_1668		N	High
SAUSA300_1231		N	High	SAUSA300_1867		N	High
SAUSA300_2366	*hlgC*	N	Low	SAUSA300_1269	*femA*	N	Low
SAUSA300_0521	*nusG*	N	Low	SAUSA300_0693		N	Low
SAUSA300_0072		N	Low	SAUSA300_2428		N	Low
SAUSA300_2551	*nrdD*	N	Low	SAUSA300_0336		N	Low
SAUSA300_1127	*smc*	N	Low	SAUSA300_0651		N	Low
SAUSA300_0849	*mnhG*	N	High	SAUSA300_2185	*rplO*	N	Low
SAUSA300_0824		N	High	SAUSA300_0106		N	High
SAUSA300_0169		N	High	SAUSA300_2275		N	Low
SAUSA300_0988	*trkA*	N	High	SAUSA300_1470		N	Low
SAUSA300_0919	*murE*	N	Low	SAUSA300_2571	*argR (3)*	N	High
SAUSA300_0816		N	Low	SAUSA300_0989		N	High
SAUSA300_1107		N	High	SAUSA300_0423		N	High
SAUSA300_2444	*gntR*	N	Low	SAUSA300_1321		N	Low
SAUSA300_0079		N	Low	SAUSA300_1520		N	High
SAUSA300_0933		N	High	SAUSA300_2042		N	High
SAUSA300_0714		N	High	SAUSA300_1982	*groEL*	N	Low
SAUSA300_1285		N	Low	SAUSA300_0823		N	High
SAUSA300_1509		N	High	SAUSA300_0227	*fadD*	N	Low
SAUSA300_2363		N	High	SAUSA300_2346	*nirB*	N	Low
SAUSA300_2406		N	High	SAUSA300_2375		N	High
SAUSA300_2376		N	High	SAUSA300_0653		N	High
SAUSA300_1512	*pbp3*	N	Low	SAUSA300_2149	*lacG*	N	Low
SAUSA300_0680	*norA*	N	High	SAUSA300_1697		N	Low
SAUSA300_1395		N	Low	SAUSA300_0792		N	High
SAUSA300_2520		N	High	SAUSA300_1890		N	Low
SAUSA300_2145		N	High	SAUSA300_2385		N	High

555-595 596-636

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SAUSA300_0420		N	Low	SAUSA300_0226		N	Low
SAUSA300_1702		N	Low	SAUSA300_1389		N	Low
SAUSA300_0711		N	High	SAUSA300_2088	*luxS*	N	Low
SAUSA300_1172		N	Low	SAUSA300_0241		N	High
SAUSA300_1301		N	Low	SAUSA300_2113		N	High
SAUSA300_0451		N	Low	SAUSA300_0631		N	High
SAUSA300_1641	*gltA*	N	Low	SAUSA300_1213		N	High
SAUSA300_0283		N	High	SAUSA300_1318		N	High
SAUSA300_0580		N	High	SAUSA300_1496		N	Low
SAUSA300_1454	*zwf*	N	Low	SAUSA300_2634		N	High
SAUSA300_1770		N	High	SAUSA300_1598	*ruvA*	N	Low
SAUSA300_0252		N	Low	SAUSA300_2490		N	High
SAUSA300_1706		N	Low	SAUSA300_0417		N	Low
SAUSA300_0854	*mnhB*	N	High	SAUSA300_1135		N	High
SAUSA300_1503		N	High	SAUSA300_1698		N	Low
SAUSA300_1197		N	Low	SAUSA300_1048	*sdhB*	N	Low
SAUSA300_1759		N	Low	SAUSA300_0625	*tagG*	N	High
SAUSA300_2284		N	Low	SAUSA300_1027	*rpmF*	N	Low
SAUSA300_0346		N	High	SAUSA300_0010		N	High
SAUSA300_1143	*topA*	N	Low	SAUSA300_1807		N	Low
SAUSA300_0745		N	High	SAUSA300_1801	*fumC*	N	Low
SAUSA300_2298		N	High	SAUSA300_1244	*mscL*	N	High
SAUSA300_0354	*ltrA*	N	High	SAUSA300_2286		N	High
SAUSA300_1601	*rpmA*	N	Low	SAUSA300_0950	*sspB*	N	Low
SAUSA300_1973		N	Low	SAUSA300_1442	*srrA*	N	Low
SAUSA300_2067	*glyA*	N	Low	SAUSA300_1957		N	Low
SAUSA300_0400		N	Low	SAUSA300_1012		N	Low
SAUSA300_2471		N	High	SAUSA300_1610	*folC*	N	High
SAUSA300_2131		N	High	SAUSA300_2272		N	Low
SAUSA300_0948	*menB*	N	Low	SAUSA300_1191	*glpF*	N	High
SAUSA300_2359		N	Low	SAUSA300_0307		N	Low
SAUSA300_2504		N	High	SAUSA300_0880		N	High
SAUSA300_1423	*polA*	N	Low	SAUSA300_2448		N	High
SAUSA300_0063		N	High	SAUSA300_0049		N	High
SAUSA300_2630	*nixA*	N	High	SAUSA300_0826		N	Low
SAUSA300_1471	*xseB*	N	Low	SAUSA300_1208		N	High
SAUSA300_2126		N	High	SAUSA300_0949	*sspC*	N	High
SAUSA300_1535	*rpsU*	N	Low	SAUSA300_1218		N	High
SAUSA300_1193	*glpD*	N	Low	SAUSA300_1046	*sdhC*	N	High
SAUSA300_1958		N	Low	SAUSA300_0310	*pfoR*	N	High
SAUSA300_1632	*nrdR*	N	Low	SAUSA300_1252		N	High

637-676 677-716

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SAUSA300_2412		N	Low	SAUSA300_1994	*scrB*	N	Low
SAUSA300_0413		N	Low	SAUSA300_1916		N	Low
SAUSA300_2644	*gidB*	N	Low	SAUSA300_0638		N	High
SAUSA300_2574		N	Low	SAUSA300_1960		N	Low
SAUSA300_0012		N	Low	SAUSA300_0279		N	Low
SAUSA300_1907		N	High	SAUSA300_1619	*hemA*	N	High
SAUSA300_0899		N	Low	SAUSA300_1876		N	Low
SAUSA300_2173	*truA*	N	Low	SAUSA300_0118		N	High
SAUSA300_2341	*narJ*	N	Low	SAUSA300_1230		N	Low
SAUSA300_1489		N	Low	SAUSA300_0788		N	Low
SAUSA300_0814		N	Low	SAUSA300_1540	*dnaK*	N	Low
SAUSA300_0981		N	Low	SAUSA300_0038	*ccrA*	N	High
SAUSA300_0256		N	High	SAUSA300_0375		N	Low
SAUSA300_2310		N	High	SAUSA300_2322		N	High
SAUSA300_1047	*sdhA*	N	Low	SAUSA300_1884		N	Low
SAUSA300_1565		N	High	SAUSA300_0843		N	Low
SAUSA300_0288		N	Low	SAUSA300_2622		N	Low
SAUSA300_0253	*scdA*	N	Low	SAUSA300_1971		N	High
SAUSA300_1146	*hslV*	N	Low	SAUSA300_2210	*glcU*	N	High
SAUSA300_0655		N	Low	SAUSA300_2282		N	High
SAUSA300_0348		N	High	SAUSA300_1243	*sbcC*	N	High
SAUSA300_0444	*gltC*	N	High	SAUSA300_1376		N	Low
SAUSA300_2269		N	High	SAUSA300_1105	*priA*	N	Low
SAUSA300_1898		N	High	SAUSA300_2072	*prfA*	N	Low
SAUSA300_1396		N	Low	SAUSA300_0539	*ilvE*	N	Low
SAUSA300_2187	*rpsE*	N	Low	SAUSA300_1915		N	High
SAUSA300_2253	*ssaA (2)*	N	Low	SAUSA300_0033		N	High
SAUSA300_1234	*rpsN*	N	Low	SAUSA300_0562	*thiD*	N	Low
SAUSA300_1750		N	High	SAUSA300_0347	*tatC*	N	High
SAUSA300_0496	*lysS*	N	Low	SAUSA300_1502		N	High
SAUSA300_1860	*pepS*	N	Low	SAUSA300_1670	*serA*	N	High
SAUSA300_0453		N	Low	SAUSA300_2165	*budA*	N	Low
SAUSA300_1350		N	Low	SAUSA300_1733		N	Low
SAUSA300_2484		N	Low	SAUSA300_0363		N	High
SAUSA300_1850		N	High	SAUSA300_1317	*msrA (2)*	N	Low
SAUSA300_2167		N	High	SAUSA300_2086		N	Low
SAUSA300_1790	*prsA*	N	Low	SAUSA300_1575	*alaS*	N	Low
SAUSA300_1956		N	Low	SAUSA300_1421		N	Low
SAUSA300_2462	*frp*	N	Low	SAUSA300_0707		N	High
SAUSA300_0221	*pflA*	N	Low	SAUSA300_0177		N	Low

717-756 757-796

Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low	Gene Locus	Gene Name	Sig. in Bonferroni?	High/Low
SAUSA300_1408		N	Low	SAUSA300_1176	*pgsA*	N	High
SAUSA300_0585		N	High	SAUSA300_2044	*cls*	N	High
SAUSA300_0409		N	Low	SAUSA300_0463		N	Low
SAUSA300_2409		N	High	SAUSA300_0629	*pbp4*	N	Low
SAUSA300_1343	*nth*	N	Low	SAUSA300_2485		N	Low
SAUSA300_1902		N	Low	SAUSA300_1769	*lukE*	N	Low
SAUSA300_0005	*gyrB*	N	Low	SAUSA300_0838	*dltD*	N	Low
SAUSA300_2076		N	Low	SAUSA300_2197	*rplP*	N	Low
SAUSA300_0054		N	High	SAUSA300_1943		N	Low
SAUSA300_0570	*eutD*	N	Low	SAUSA300_1522	*dnaG*	N	Low
SAUSA300_0099	*plc*	N	Low	SAUSA300_2179	*rpsK*	N	Low
SAUSA300_1006		N	Low	SAUSA300_1803		N	Low
SAUSA300_2261		N	High	SAUSA300_0378		N	High
SAUSA300_0179		N	Low	SAUSA300_0716		N	Low
SAUSA300_0084		N	High	SAUSA300_1546	*holA*	N	High
SAUSA300_0762	*secG*	N	High	SAUSA300_2229	*modB*	N	High
SAUSA300_0533	*tuf*	N	Low	SAUSA300_1337		N	Low
SAUSA300_1284		N	Low	SAUSA300_2383		N	High
SAUSA300_1354		N	Low	SAUSA300_0720		N	High
SAUSA300_2228	*modC*	N	High	SAUSA300_1665		N	High
SAUSA300_1885	*ligA*	N	Low	SAUSA300_0257		N	High
SAUSA300_1064		N	High	SAUSA300_1184		N	High
SAUSA300_0528	*ropC*	N	Low	SAUSA300_1666	*rpsD*	N	Low
SAUSA300_0305		N	High	SAUSA300_1934		N	Low
SAUSA300_0774	*empbp*	N	Low	SAUSA300_1785		N	High
SAUSA300_2505		N	High	SAUSA300_0909		N	High
SAUSA300_1977		N	High	SAUSA300_2569	*arcB (2)*	N	Low
SAUSA300_0355		N	Low	SAUSA300_1792		N	Low
SAUSA300_2613	*hisZ*	N	Low	SAUSA300_2137		N	High
SAUSA300_0587		N	High	SAUSA300_2063	*atpE*	N	High
SAUSA300_0396	*set7*	N	High	SAUSA300_2542		N	Low
SAUSA300_1764	*epiD*	N	High	SAUSA300_0511	*radA*	N	Low
SAUSA300_0811		N	Low	SAUSA300_2268		N	High
SAUSA300_1463		N	Low	SAUSA300_1065		N	High
SAUSA300_1766	*epiB*	N	High	SAUSA300_1281	*pstA*	N	High
SAUSA300_1210		N	High	SAUSA300_1611	*valS*	N	Low
SAUSA300_1377		N	Low	SAUSA300_0575		N	Low
SAUSA300_2329	*gltT*	N	High	SAUSA300_2500		N	Low
SAUSA300_2608	*hisH*	N	High	SAUSA300_1295		N	Low
SAUSA300_0092		N	High	SAUSA300_1425		N	Low

797-828

SAUSA300_1511	*rpmG (2)*	N	Low
SAUSA300_2354		N	Low
SAUSA300_0465		N	Low
SAUSA300_0273		N	High
SAUSA300_0538		N	Low
SAUSA300_0746		N	Low
SAUSA300_1789		N	High
SAUSA300_0132		N	High
SAUSA300_0649		N	High
SAUSA300_2530		N	High
SAUSA300_0514	*cysE*	N	High
SAUSA300_2184	*secY*	N	High
SAUSA300_0438		N	Low
SAUSA300_2009	*ilvC*	N	Low
SAUSA300_0822	*sufB*	N	Low
SAUSA300_0174		N	High
SAUSA300_0490	*hslO*	N	Low
SAUSA300_1150	*tsf*	N	Low
SAUSA300_1403		N	Low
SAUSA300_2277	*hutI*	N	Low
SAUSA300_0255		N	High
SAUSA300_0160	*cap5I*	N	High
SAUSA300_2075	*rho*	N	Low
SAUSA300_2487	*feoB*	N	High
SAUSA300_1581		N	Low
SAUSA300_1312		N	Low
SAUSA300_2457		N	Low
SAUSA300_1058		N	Low
SAUSA300_1550		N	Low
SAUSA300_1525	*glyS*	N	Low
SAUSA300_0887	*oppB*	N	High
SAUSA300_1495		N	High

Appendix B

References

1. Lowy, F. D. Antimicrobial Resistance: the Example of *Staphylococcus aureus*. *J. Clin. Invest.* **111,** 1265–1273 (2003).
2. Chambers, H. F. & DeLeo, F. R. Waves of Resistance: *Staphylococcus aureus* in the Antibiotic Era. *Nat. Rev. Microbiol.* **7,** 629–641 (2009).
3. Tong, S. Y. . C., Chen, L. F. & Fowler Jr, V. G. Colonization, Pathogenicity, Host Susceptibility and Therapeutics for *Staphylococcus aureus*: What is the Clinical Relevance? *Semin. Immunopathol.* **34,** 185–200 (2012).
4. Gordon, R. J. & Lowy, F. D. Pathogenesis of Methicillin-Resistant *Staphylococcus aureus* Infection. *Clin. Infect. Dis.* **46,** 350–359 (2008).
5. Wertheim, H. F. L. *et al.* The Role of Nasal Carriage in *Staphylococcus aureus* Infections. *Lancet Infect. Dis.* **5,** 751–762 (2005).
6. Archer, G. L. *Staphylococcus aureus*: A Well-Armed Pathogen. *Clin. Infect. Dis.* **26,** 1179–1181 (1998).
7. Rasigade, J.-P. & Vandenesch, F. *Staphylococcus aureus*: a Pathogen with Still Unresolved Issues. *Infect. Genet. Evol.* **21,** 510–514 (2014).
8. Sibbald, M. J. J. B. *et al.* Mapping the Pathways to Staphylococcal Pathogenesis by Comparative Secretomics. *Microbiol. Mol. Biol. Rev.* **70,** 755–788 (2006).
9. Young, B. C. *et al.* Evolutionary Dynamics of *Staphylococcus aureus* During Progression from Carriage to Disease. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 4550–4555 (2012).
10. Simor, A. E. Staphylococcal Decolonisation: an Effective Strategy for Prevention of Infection? *Lancet Infect. Dis.* **11,** 952–962 (2011).
11. Nouwen, J., Boelens, H., van Belkum, A. & Verburgh, H. Human Factor in *Staphylococcus aureus* Nasal Carriage. *Infect. Immun.* **72,** 6685–6688 (2004).
12. van Belkum, A. *et al.* Reclassification of *Staphylococcus aureus* Nasal Carriage Types. *J. Infect. Dis.* **199,** 1820–1826 (2009).
13. Hu, L., Umeda, A., Kondo, S. & Amako, K. Typing of *Staphylococcus aureus* Colonising Human Nasal Carriers by Pulsed-Field Gel Electrophoresis. *J. Med. Microbiol.* **42,** 127–132 (1995).
14. Riewerts Eriksen, N. H., Espersen, F., Thamdrup Rosdahl, V. & Jensen, K. Carriage of *Staphylococcus aureus* Among 104 Healthy Persons During a 19-month Period. *Epidemiol. Infect.* **115,** 51–60 (1995).

15. Vandenbergh, M. F. Q. *et al.* Follow-Up of *Staphylococcus aureus* Nasal Carriage after 8 Years : Redefining the Persistent Carrier State. *J. Clin. Microbiol.* **37,** 3133–3140 (1999).

16. Cole, A. M. *et al.* Determinants of *Staphylococcus aureus* Nasal Carriage. *Clin. Diagn. Lab. Immunol.* **8,** 1064–1069 (2001).

17. Lipsky, B. A., Pecoraro, R. E., Chen, M. S. & Koepsell, T. D. Factors Affecting Staphylococcal Colonization Among NIDDM Outpatients. *Diabetes Care* **10,** 483 LP-486 (1987).

18. Peacock, S. J., de Silva, I. & Lowy, F. D. What Determines Nasal Carriage of *Staphylococcus aureus*? *Trends Microbiol.* **9,** 605–610 (2001).

19. Nouwen, J. L. *et al.* Predicting the *Staphylococcus aureus* Nasal Carrier State: Derivation and Validation of a 'Culture Rule'. *Clin. Infect. Dis.* **39,** 806–811 (2004).

20. Candi, E., Schmidt, R. & Melino, G. The Cornified Envelope: a model of Cell Death in the Skin. *Nature Reviews Molecular Cell Biology* **6,** 328–340 (2005).

21. Foster, T. J., Geoghegan, J. A., Ganesh, V. K. & Höök, M. Adhesion, Invasion and Evasion: the Many Functions of the Surface Proteins of *Staphylococcus aureus*. *Nat. Rev. Microbiol.* **12,** 49–62 (2013).

22. Lacey, K. A., Geoghegan, J. A. & McLoughlin, R. M. The Role of *Staphylococcus aureus* Virulence Factors in Skin Infection and Their Potential as Vaccine Antigens. *Pathogens* **5,** (2016).

23. Reddy, P. N., Srirama, K. & Dirisala, V. R. An Update on Clinical Burden, Diagnostic Tools, and Therapeutic Options of *Staphylococcus aureus*. *Infect. Dis. Res. Treat.* **10,** 1–15 (2017).

24. Mulcahy, M. E. & McLoughlin, R. M. Host-Bacterial Crosstalk Determines *Staphylococcus aureus* Nasal Colonization. *Trends Microbiol.* **24,** 872–886 (2016).

25. Son, E. D. *et al. Staphylococcus aureus* Inhibits Terminal Differentiation of Normal Human Keratinocytes by Stimulating Interleukin-6 Secretion. *J. Dermatol. Sci.* **74,** 64–71 (2014).

26. Boniface, K. *et al.* IL-22 Inhibits Epidermal Differentiation and Induces Proinflammatory Gene Expression and Migration of Human Keratinocytes. *J. Immunol.* **174,** 3695–3702 (2005).

27. Mulcahy, M. E., Leech, J. M., Renauld, J. C., Mills, K. H. G. & McLoughlin, R. M. Interleukin-22 Regulates Antimicrobial Peptide Expression and Keratinocyte Differentiation to Control *Staphylococcus aureus* Colonization of the Nasal Mucosa. *Mucosal Immunol.* **9,** 1429–1441 (2016).

28. Nagaoka, I., Hirota, S., Yomogida, S., Ohwada, A. & Hirata, M. Synergistic Actions of Antibacterial Neutrophil Defensins and Cathelicidins. *Inflamm. Res.* **49,** 73–79 (2000).

29. Ong, P. Y. *et al.* Endogenous Antimicrobial Peptides and Skin Infections in Atopic Dermatitis. *N. Engl. J. Med.* **347,** 1151–1160 (2002).

30. Cole, A. M., Dewan, P. & Ganz, T. Innate Antimicrobial Activity of Nasal Secretions Innate Antimicrobial Activity of Nasal Secretions. *Infect. Immun.* **67,** 3267–3275 (1999).

31. von Aulock, S. *et al.* Lipoteichoic Acid from *Staphylococcus aureus* is a Potent Stimulus for Neutrophil Recruitment. *Immunobiology* **208,** 413–22 (2003).

32. Colque-Navarro, P., Jacobsson, G., Andersson, R., Flock, J.-I. & Möllby, R. Levels of Antibody Against 11 *Staphylococcus aureus* Antigens in a Healthy Population. *Clin. Vaccine Immunol.* **17,** 1117–1123 (2010).

33. Verkaik, N. J. *et al.* Anti-Staphylococcal Humoral Immune Response in Persistent Nasal Carriers and Noncarriers of *Staphylococcus aureus*. *J. Infect. Dis.* **199,** 625–632 (2009).

34. Wertheim, H. F. L. *et al.* Risk and Outcome of Nosocomial *Staphylococcus aureus* Bacteraemia in Nasal Carriers Versus Non-Carriers. *Lancet* **364,** 703–705 (2004).

35. Holtfreter, S. *et al. Staphylococcus aureus* Carriers Neutralize Superantigens by Antibodies Specific for Their Colonizing Strain: a Potential Explanation for Their Improved Prognosis in Severe Sepsis. *J.Infect.Dis.* **193,** 1275–1278 (2006).

36. Kolata, J. *et al.* Distinctive Patterns in the Human Antibody Response to *Staphylococcus aureus* Bacteremia in Carriers and Non-Carriers. *Proteomics* **11,** 3914–3927 (2011).

37. Holtfreter, S. *et al.* Human Immune Proteome in Experimental Colonization with *Staphylococcus aureus*. *Clin. Vaccine Immunol.* **16,** 1607–1614 (2009).

38. Verkaik, N. J. *et al.* Induction of Antibodies by *Staphylococcus aureus* Nasal Colonization in Young Children. *Clin. Microbiol. Infect.* **16,** 1312–1317 (2010).

39. Cole, A. L. *et al.* Host Innate Inflammatory Factors and Staphylococcal Protein A Influence the Duration of Human *Staphylococcus aureus* Nasal Carriage. *Mucosal Immunol.* **9,** 1538–1548 (2016).

40. Uehara, Y. *et al.* Bacterial Interference Among Nasal Inhabitants: Eradication of Staphylococcus aureus from Nasal Cavities by Artificial Implantation of Corynebacterium sp. *J. Hosp. Infect.* **44,** 127–133 (2000).

41. Bogaert, D. *et al.* Colonisation by *Streptococcus pneumoniae* and *Staphylococcus aureus* in Healthy Children. *Lancet* **363,** 1871–1872 (2004).

42. Frank, D. N. *et al.* The Human Nasal Microbiota and *Staphylococcus aureus* Carriage. *PLoS One* **5,** e10598 (2010).

43. van Gils, E. J. M. *et al.* Effect of Seven-Valent Pneumococcal Conjugate Vaccine on *Staphylococcus aureus* Colonisation in a Randomised Controlled Trial. *PLoS One* **6,** (2011).

44. Selva, L. *et al.* Killing Niche Competitors by Remote-Control Bacteriophage Induction. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 1234–8 (2009).

45. Ramsey, M. M., Freire, M. O., Gabrilska, R. A., Rumbaugh, K. P. & Lemon, K. P. *Staphylococcus aureus* Shifts toward commensalism in response to Corynebacterium species. *Front. Microbiol.* **7,** 1–15 (2016).

46. Sonesson, A. *et al.* Identification of Bacterial Biofilm and the *Staphylococcus aureus* Derived Protease, Staphopain, on the Skin Surface of Patients with Atopic Ddermatitis. *Sci. Rep.* **7,** 8689 (2017).

47. Hepburn, L. *et al.* The Complex Biology and Contribution of *Staphylococcus aureus* in Atopic Dermatitis, Current and Future Therapies. *Br. J. Dermatol.* **177,** 63–71 (2017).

48. Lin, Y. T., Wang, C. T. & Chiang, B. L. Role of Bacterial Pathogens in Atopic Dermatitis. *Clin Rev Allergy Immunol* **33,** 167–177 (2007).

49. Kong, H. H. *et al.* Temporal Shifts in the Skin Microbiome Associated with Disease Flare and Treatment in Children with Atopic Dermatitis. *Genome Res.* **22,** 850–859 (2012).

50. Pournaras, C. C., Lübbe, J. & Saurat, J. H. Staphylococcal Colonization in Atopic Dermatitis Treatment with Topical Tacrolimus (Fk506). *J. Invest. Dermatol.* **116,** 480–481 (2001).

51. Foster, T. J. & Höök, M. Surface Protein Adhesins of *Staphylococcus aureus*. *Trends Microbiol.* **6,** 484–488 (1998).

52. Cho, S. H., Strickland, I., Boguniewicz, M. & Leung, D. Y. M. Fibronectin and Fibrinogen Contribute to the Enhanced Binding of *Staphylococcus aureus* to Atopic Skin. *J. Allergy Clin. Immunol.* **108,** 269–274 (2001).

53. Totté, J. *et al.* Targeted Anti-Staphylococcal Therapy with Endolysins in Atopic Dermatitis and the Effect on Steroid Use, Disease Severity and the Microbiome: Study Protocol for a Randomized Controlled Trial (MAAS trial). *Trials* **18,** 404 (2017).

54. Bran, J. L., Levison, M. E. & Kaye, D. Survey for Methicillin-Resistant Staphylococci. *Antimicrob. Agents Chemother.* **1,** 235–236 (1972).

55. Crisostomo, M. I. *et al.* The Evolution of Methicillin Resistance in *Staphylococcus aureus*: Similarity of Genetic Backgrounds in Historically Early Methicillin-Susceptible and -Resistant Isolates and Contemporary Epidemic Clones. *Proc. Natl. Acad. Sci.* **98,** 9865–9870 (2001).

56. Enright, M. C. *et al.* The Evolutionary History of Methicillin-Resistant *Staphylococcus aureus* (MRSA). *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7687–7692 (2002).

57. Hiramatsu, K. *et al.* Dissemination in Japanese Hospitals of Strains of *Staphylococcus aureus* Heterogeneously Resistant to Vancomycin. *Lancet* **350,** 1670–1673 (1997).

58. Weigel, L. M. *et al.* Genetic Analysis of a High-Level Vancomycin-Resistant Isolate of *Staphylococcus aureus*. *Science (80-.).* **302,** 1569–1571 (2003).

59. Centers for Disease Control and Prevention (CDC). Four Pediatric Deaths from Community-Acquired Methicillin-Resistant *Staphylococcus aureus* — Minnesota and North Dakota, 1997-1999. *Morb. Mortal. Wkly. Rep.* **48,** 707—710 (1999).

60. O'Brien, F. G. *et al.* Diversity Among Community Isolates of Methicillin-Reistant *Staphylococcus aureus* Strains Causing Outpaitent Infections in Australia. *J Clin Microbiol* **42,** 4735–43 (2004).

61. Genestier, A.-L. *et al. Staphylococcus aureus* Panton-Valentine Leukocidin Directly Targets Mitochondria and Induces Bax-Independent Apoptosis of Human Neutrophils. *J. Clin. Invest.* **115,** 3117–3127 (2005).

62. Tenover, F. C. & Goering, R. V. Methicillin-Resistant *Staphylococcus aureus* Strain USA300: Origin and Epidemiology. *J. Antimicrob. Chemother.* **64,** 441–446 (2009).

63. Seybold, U. *et al.* Emergence of Community-Associated Methicillin-Resistant *Staphylococcus aureus* USA300 Genotype as a Major Cause of Health Care-Associated Blood Stream Infections. *Clin. Infect. Dis.* **42,** 647–656 (2006).

64. David, M. Z. & Daum, R. S. Community-Associated Methicillin-Resistant *Staphylococcus aureus*: Epidemiology and Clinical Consequences of an Emerging Epidemic. *Clin. Microbiol. Rev.* **23,** 616–687 (2010).

65. Rudkin, J. K. *et al.* Methicillin Resistance Reduces the Virulence of Healthcare-Associated Methicillin-Resistant *Staphylococcus aureus* by Interfering With the *agr* Quorum Sensing System. *J. Infect. Dis.* **205,** 798–806 (2012).

66. Chen, C.-J., Wang, S.-C., Chang, H.-Y. & Huang, Y.-C. Longitudinal Analysis of Methicillin-Resistant and Methicillin-Susceptible *Staphylococcus aureus* Carriage in Healthy Adolescents. *J. Clin. Microbiol.* **51,** 2508–2514 (2013).

67. Dulon, M., Peters, C., Schablon, A. & Nienhaus, A. MRSA Carriage Among Healthcare Workers in Non-Outbreak Settings in Europe and the United States: a Systematic Review. *BMC Infect. Dis.* **14,** (2014).

68. Liu, W. *et al.* The Prevalence and Influencing Factors of Methicillin-Resistant *Staphylococcus aureus* Carriage in People in Contact with Livestock: a Systematic Review. *Am. J. Infect. Control* **43,** 469–475 (2015).

69. McKinnell, J. A., Miller, L. G., Eells, S. J., Cui, E. & Huang, S. S. A Systematic Literature Review and Meta-Analysis of Factors Associated with Methicillin-Resistant Staphylococcus aureus Colonization at Time of Hospital or Intensive Care Unit Admission. *Infect. Control Hosp. Epidemiol.* **34,** (2013).

70. Weinstein, R. A., Milstone, A. M., Passaretti, C. L. & Perl, T. M. Chlorhexidine: Expanding the Armamentarium for Infection Control and Prevention. *Clin. Infect. Dis.* **46,** 274–281 (2008).

71. Jones, C. G. Chlorhexidine: Is it still the gold standard? *Periodontol. 2000* **15,** 55–62 (1997).

72. Paulsen, I. T. *et al.* The SMR Family: a Novel Family of Multidrug Efflux Proteins Involved with the Efflux of Lipophilic Drugs. *Mol. Microbiol.* **19,** 1167–1175 (1996).

73. Paulsen, I. T., Brown, M. H., Littlejohn, T. G., Mitchell, B. A. & Skurray, R. A. Multidrug Resistance Proteins QacA and QacB from *Staphylococcus aureus*: Membrane Topology and Identification of Residues Involved in Substrate Specificity. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 3630–5 (1996).

74. Schlett, C. D. *et al.* Prevalence of Chlorhexidine-Resistant Methicillin-Resistant *Staphylococcus aureus* Following Prolonged Exposure. *Antimicrob. Agents Chemother.* **58,** 4404–4410 (2014).

75. Thomas, C. M., Hothersall, J., Willis, C. L. & Simpson, T. J. Resistance To and Synthesis of the Antibiotic Mupirocin. *Nat. Rev. Microbiol.* **8,** 281–289 (2010).

76. Hodgson, J. E. *et al.* Molecular Characterization of the Gene Encoding High-Level Mupirocin Resistance in *Staphylococcus aureus* J2870. *Antimicrob. Agents Chemother.* **38,** 1205–8 (1994).

77. Seah, C. *et al.* MupB, a New High-Level Mupirocin Resistance Mechanism in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **56,** 1916–1920 (2012).

78. Gilbart, J., Perry, C. R. & Slocombe, B. High-Level Mupirocin Resistance in *Staphylococcus aureus*: Evidence for Two Distinct Isoleucyl-tRNA Synthetases. *Antimicrob. Agents Chemother.* **37,** 32–38 (1993).

79. Hurdle, J. G., O'Neil, A. J., Mody, L., Chopra, I. & Bradley, S. F. *In vivo* Transfer of High-Level Mupirocin Resistance from *Staphylococcus epidermidis* to Methicillin-Resistant *Staphylococcus aureus* Associated with Failure of Mupirocin Prophylaxis. *J Antimicrob Chemother* **56,** 1166–1168 (2005).

80. Needham, C., Rahman, M., Dyke, K. G. & Noble, W. C. An Investigation of Plasmids from *Staphylococcus aureus* that Mediate Resistance to Mupirocin and Tetracycline. *Microbiology* **140,** 2577–83 (1994).

81. Ramsey, M. A., Bradley, S. F., Kauffman, C. A. & Morton, T. M. Identification of Chromosomal Location of *mupA* Gene, Encoding Low-Level Mupirocin Resistance in Staphylococcal Isolates. *Antimicrob. Agents Chemother.* **40,** 2820–3 (1996).

82. Fujimura, S., Watanabe, A. & Beighton, D. Characterization of the *mupA* Gene in Strains of Methicillin-Resistant *Staphylococcus aureus* with a Low Level of Resistance to Mupirocin. *Antimicrob. Agents Chemother.* **45,** 641–642 (2001).

83. Udo, E. E., Al-Sweih, N. & Noronha, B. C. A Chromosomal Location of the *mupA* Gene in *Staphylococcus aureus* Expressing High-Level Mupirocin Resistance. *J. Antimicrob. Chemother.* **51,** 1283–1286 (2003).

84. Driscoll, D. G., Young, C. L. & Ochsner, U. A. Transient Loss of High-Level Mupirocin Resistance in Staphylococcus aureus Due to MupA Polymorphism. *Antimicrob. Agents Chemother.* **51,** 2247–2248 (2007).

85. GlaxoSmithKline. Package Leaflet: Information for the user Bactroban® 2% Nasal Ointment (mupirocin calcium).

86. Watanabe, H. *et al.* Emergence and Spread of Low-Level Mupirocin Resistance in Methicillin-Resistant *Staphylococcus aureus* Isolated from a Community Hospital in Japan. *J. Hosp. Infect.* **47,** 294–300 (2001).

87. Cookson, B. D. The emergence of mupirocin resistance: A challenge to infection control and antibiotic prescribing practice. *J. Antimicrob. Chemother.* **41,** 11–18 (1998).

88. Semret, M. & Miller, M. A. Topical Mupirocin for Eradication of MRSA Colonization with Mupirocin-Resistant Strains. *Infect. Control Hosp. Epidemiol.* **22,** 578–580 (2001).

89. Hetem, D. J. & Bonten, M. J. M. Clinical Relevance of Mupirocin Resistance in *Staphylococcus aureus*. *J. Hosp. Infect.* **85,** 249–256 (2013).

90. Walker, E. S., Vasquez, J. E., Dula, R., Bullock, H. & Sarubbi, F. A. Mupirocin-Resistant, Methicillin-Resistant *Staphylococcus aureus*: Does Mupirocin Remain Effective? *Infect. Control Hosp. Epidemiol.* **24,** 342–346 (2003).

91. Lee, A. S. *et al.* Impact of Combined Low-Level Mupirocin and Genotypic Chlorhexidine Resistance on Persistent Methicillin-Resistant *Staphylococcus aureus* Carriage After Decolonization Therapy: A Case-Control Study. *Clin. Infect. Dis.* **52,** 1422–1430 (2011).

92. Sai, N., Laurent, C., Strale, H., Denis, O. & Byl, B. Efficacy of the Decolonization of Methicillin-Resistant *Staphylococcus aureus* Carriers in Clinical Practice. *Antimicrob. Resist. Infect. Control* **4,** 1–8 (2015).

93. Hanaki, H. *et al.* Activated Cell-Wall Synthesis is Associated with Vancomycin Resistance in Methicillin-Resistant *Staphylococcus aureus* Clinical strains Mu3 and Mu50. *J. Antimicrob. Chemother.* **42,** 199–209 (1998).

94. Showsh, S. A., De Boever, E. H. & Clewell, D. B. Vancomycin Resistance Plasmid in *Enterococcus faecalis* That Encodes Sensitivity to a Sex Pheromone Also Produced by *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **45,** 2177–2179 (2001).

95. Hooper, D. C. Fluoroquinolone Resistance Among Gram-Positive Cocci. *Lancet Infect. Dis.* **2,** 530–538 (2002).

96. Cheung, A. L. & Fischetti, V. A. The Role of Fibrinogen in Mediating Staphylococcal Adherence to Fibers. *J. Surg. Res.* **50,** 150–155 (1991).

97. Vaudaux, P. *et al.* Fibronectin is More Active than Fibrin or Fibrinogen in Promoting *Staphylococcus aureus* Adherence to Inserted Intravascular Catheters. *J. Infect. Dis.* **167,** 633–641 (1993).

98. Lowy, F. D. *Staphylococcus aureus* Infections. *N. Engl. J. Med.* **339,** 520–532 (1998).

99. Tong, S. Y. C., Davis, J. S., Eichenberger, E., Holland, T. L. & Fowler, V. G. *Staphylococcus aureus* Infections: Epidemiology, Pathophysiology, Clinical Manifestations, and Management. *Clin. Microbiol. Rev.* **28,** 603–661 (2015).

100. Fridkin, S. K. *et al.* Methicillin-Resistant *Staphylococcus aureus* Disease in Three Communities. *N. Engl. J. Med.* **352,** 1436–1444 (2005).

101. Kaplan, S. L. *et al.* Three-Year Surveillance of Community-Acquired *Staphylococcus aureus* Infections in Children. *Clin.Infect.Dis.* **40,** 1785–1791 (2005).

102. Purcell, K. & Fergie, J. Epidemic of Community-Acquired Methicillin-Resistant *Staphylococcus aureus* Infections: a 14-Year Study at Driscoll Children's Hospital. *Arch Pediatr Adolesc Med* **159,** 980–985 (2005).

103. Davis, S. L. *et al.* Epidemiology and Outcomes of Community-Associated Methicillin-Resistant *Staphylococcus aureus* Infection. *J. Clin. Microbiol.* **45,** 1705–1711 (2007).

104. Liu, C. *et al.* A Population-Based Study of the Incidence and Molecular Epidemiology of Methicillin-Resistant *Staphylococcus aureus* Disease in San Francisco, 2004–2005. *Clin. Infect. Dis.* **46,** 1637–1646 (2008).

105. Tong, S. Y. C. *et al.* Community Associated Strains of Methicillin-Resistant *Staphylococcus aureus* and Methicillin-Susceptible *S. aureus* in Indigenous Northern Australia: Epidemiology and Outcomes. *J. Infect. Dis.* **199,** 1461–1470 (2009).

106. Bangert, S., Levy, M. & Hebert, A. A. Bacterial Resistance and Impetigo Treatment Trends: A Review. *Pediatr. Dermatol.* **29,** 243–248 (2012).

107. Fowler, V. G. *et al. Staphylococcus aureus* Bacteremia After Median Sternotomy: Clinical Utility of Blood Culture Results in the Identification of Postoperative Mediastinitis. *Circulation* **108,** 73–78 (2003).

108. Ibler, K. S. & Kromann, C. B. Recurrent Furunculosis - Challenges and Management: a Review. *Clin. Cosmet. Investig. Dermatol.* **7,** 59–64 (2014).

109. Kramer, A., Schwebke, I. & Kampf, G. How Long Do Nosocomial Pathogens Persist on Inanimate Surfaces? A Systematic Review. *BMC Infect. Dis.* **6,** 130 (2006).

110. Lina, G. *et al.* Involvement of Panton-Valentine Leukocidin-Producing *Staphylococcus aureus* in Primary Skin Infections and Pneumonia. *Clin. Infect. Dis.* **29,** 1128–1132 (1999).

111. Uhlén, M. *et al.* Complete Sequence of the Staphylococcal Gene Encoding Protein A - a Gene Evolved Through Multiple Duplications. *J. Biol. Chem.* **259,** 1695–1702 (1984).

112. Hassoun, A., Linden, P. K. & Friedman, B. Incidence, Prevalence, and Management of MRSA Bacteremia Across Patient Populations—a Review of Recent Developments in MRSA Management and Treatment. *Crit. Care* **21,** 211 (2017).

113. van Hal, S. J. *et al.* Predictors of Mortality in *Staphylococcus aureus* Bacteremia. *Clin. Microbiol. Rev.* **25,** 362–386 (2012).

114. Edwards, J. R., Horan, T. C., Jernigan, J. A. & Fridkin, S. K. Central Line – Associated Bloodstream Infections in US Intensive Care Units , 1997-2007. *JAMA* **301,** 727–736 (2009).

115. FitzGerald, S. F. *et al.* A 12-year Review of *Staphylococcus aureus* Bloodstream Infections in Haemodialysis Patients: More Work to be Done. *J. Hosp. Infect.* **79,** 218–221 (2011).

116. Suzuki, M. *et al.* Bacteremia in Hemodialysis Patients. *World J. Nephrol.* **5,** 489 (2016).

117. Raad, I. *et al.* Central Venous Catheters Coated with Minocycline and Rifampin for the Prevention of Catheter-Related Colonization and Bloodstream Infections: a Randomized, Double-Blind Trial. *Ann. Intern. Med.* **127,** 267–274 (1997).

118. Chatzinikolaou, I. *et al.* Antibiotic-Coated Hemodialysis Catheters for the Prevention of Vascular Catheter-Related Infections: a Prospective, Randomized Study. *Am. J. Med.* **115,** 352–357 (2003).

119. Bor, D. H., Woolhandler, S., Nardin, R., Brusch, J. & Himmelstein, D. U. Infective Endocarditis in the U.S., 1998-2009: A Nationwide Study. *PLoS One* **8,** 1–8 (2013).

120. Holland, T. L. *et al.* Infective Endocarditis. *Nat. Rev. Dis. Prim.* **2,** 16059 (2016).

121. Patti, J. M., Allen, B. L., McGavin, M. J. & Höök, M. MSCRAMM-Mediated Adherence of Microorganisms to Host Tissues. *Annu. Rev. Microbiol.* **48,** 585–617 (1994).

122. Wolff, M., Witchitz, S., Chastang, C., Régnier, B. & Vachon, F. Prosthetic Valve Endocarditis in the ICU: Prognostic Factors of Overall Survival in a Series of 122 Cases and Consequences for Treatment Decision. *Chest* **108,** 688–694 (1995).

123. Wang, A. *et al.* Contemporary Clinical Profile and Outcome of Prosthetic Valve Endocarditis. *JAMA* **297,** 1354–1361 (2007).

124. El-Ahdab, F. *et al.* Risk of Endocarditis Among Patients with Prosthetic Valves and *Staphylococcus aureus* Bacteremia. *Am. J. Med.* **118,** 225–229 (2005).

125. Josefsson, E., Hartford, O., O'Brien, L., Patti, J. M. & Foster, T. Protection against Experimental *Staphylococcus aureus* Arthritis by Vaccination with Clumping Factor A, a Novel Virulence Determinant. *J. Infect. Dis.* **184,** 1572–1580 (2001).

126. Xiong, Y. Q. *et al.* Endovascular Infections Caused by Methicillin-Resistant *Staphylococcus aureus* are Linked to Clonal Complex-Specific Alterations in Binding and Invasion Domains of Fibronectin-Binding Protein A As Well As the Occurrence of *fnbB*. *Infect. Immun.* **83,** 4772–4780 (2015).

127. Arrecubieta, C. *et al.* The Role of Staphylococcus aureus Adhesins in the Pathogenesis of Ventricular Assist Device–Related Infections. *J. Infect. Dis.* **193,** 1109–1119 (2006).

128. Vergara-Irigaray, M. *et al.* Relevant Role of Fibronectin-Binding Proteins in *Staphylococcus aureus* Biofilm-Associated Foreign-Body Infections. *Infect. Immun.* **77,** 3978–3991 (2009).

129. Moormeier, D. E. & Bayles, K. W. *Staphylococcus aureus* Biofilm: a Complex Developmental Organism. *Mol. Microbiol.* **104,** 365–376 (2017).

130. Schneewind, O., Fowler, A. & Faull, K. F. Structure of the Cell Wall Anchor of Surface Proteins in *Staphylococcus aureus*. *Science (80-.).* **268,** 103–106 (1995).

131. McDevitt, D., Francois, P., Vaudaux, P. & Foster, T. J. Molecular Characterization of the Clumping Factor (Fibrinogen Receptor) of *Staphylococcus aureus*. *Mol. Microbiol.* **11,** 237–248 (1994).

132. Ní Eidhin, D. *et al.* Clumping factor B (ClfB), a New Surface-Located Fibrinogen-Binding Adhesin of *Staphylococcus aureus*. *Mol. Microbiol.* **30,** 245–257 (1998).

133. Hartford, O., Francois, P., Vaudaux, P. & Foster, T. J. The Dipeptide Repeat Region of the Fibrinogen-Binding Protein (Clumping Factor) is Required for Functional Expression of the Fibrinogen-Binding Domain on the *Staphylococcus aureus* Cell Surface. *Mol. Microbiol.* **25,** 1065–1076 (1997).

134. Jönsson, K., Signäs, C., Müller, H.-P. & Lindberg, M. Two Different Genes Encode Fibronectin Binding proteins in *Staphylococcus aureus* - the Complete Nucleotide Sequence and Characterization of the Second Gene. *Eur. J. Biochem.* **202,** 1041–1048 (1991).

135. Vaudaux, P. E. *et al.* Use of Adhesion Defective Mutants of *Staphylococcus aureus* to Define the Role of Specific Plasma Proteins in Promoting Adhesion to Canine Arteriovenous Shunts. *Infect. Immun.* **63,** 585–590 (1995).

136. Greene, C. *et al.* Adhesion Properties of Mutants of *Staphylococcus aureus* Defective in Fibronectin-Binding Proteins and Studies on the Expression of *fnb* Genes. *Mol. Microbiol.* **17,** 1143–1152 (1995).

137. House-Pompeo, K., Xu, Y., Joh, D., Speziale, P. & Höök, M. Conformational Changes in the Fibronectin Binding MSCRAMMs are Induced by Ligand Binding. *J. Biol. Chem.* **271,** 1379–1384 (1996).

138. Switalski, L. M. *et al.* A collagen Receptor on *Staphylococcus aureus* Strains Isolated from Patients with Septic Arthritis Mediates Adhesion to Cartilage. *Mol. Microbiol.* **7,** 99–107 (1993).

139. Ryding, U., Flock, J. I., Flock, M., Söderquist, B. & Christensson, B. Expression of Collagen-Binding Protein and Types 5 and 8 Capsular Polysaccharide in Clinical Isolates of *Staphylococcus aureus*. *J. Infect. Dis.* **176,** 1096–9 (1997).

140. Foster, T. J. Immune Evasion by Staphylococci. *Nat. Rev. Microbiol.* **3,** 948–958 (2005).

141. Ricklin, D., Hajishengallis, G., Yang, K. & Lambris, J. D. Complement: a Key System for Immune Surveillance and Homeostasis. *Nat. Immunol.* **11,** 785–797 (2010).

142. Rooijakkers, S. H. M. *et al.* Immune Evasion by a Staphylococcal Complement Inhibitor That Acts on C3 Convertases. *Nat. Immunol.* **6,** 920–927 (2005).

143. Lee, L. Y. L., Liang, X., Höök, M. & Brown, E. L. Identification and Characterization of the C3 Binding Domain of the *Staphylococcus aureus* Extracellular Fibrinogen-Binding Protein (Efb). *J. Biol. Chem.* **279,** 50710–50716 (2004).

144. Rooijakkers, S. H. M., Van Wamel, W. J. B., Ruyken, M., Van Kessel, K. P. M. & Van Strijp, J. A. G. Anti-Opsonic Properties of Staphylokinase. *Microbes Infect.* **7,** 476–484 (2005).

145. de Haas, C. J. C. *et al.* Chemotaxis Inhibitory Protein of *Staphylococcus aureus*, a Bacterial Antiinflammatory Agent. *J. Exp. Med.* **199,** 687–695 (2004).

146. Chavakis, T. *et al. Staphylococcus aureus* Extracellular Adherence Protein Serves as Anti-Inflammatory Factor by Inhibiting the Recruitment of Host Leukocytes. *Nat Med* **8,** 687–693 (2002).

147. Falugi, F., Kim, H. K. & Missiakas, D. M. Role of Protein A in the Evasion of Host Adaptive Immune Responses. *MBio* **4,** 1–9 (2013).

148. Pauli, N. T. *et al. Staphylococcus aureus* Infection Induces Protein A-Mediated Immune Evasion in Humans. *J. Exp. Med.* **211,** 2331–9 (2014).

149. Potter, K. N., Li, Y. & Capra, J. D. Staphylococcal Protein A Simultaneously Interacts with Framework Region 1, Complementarity-Determining Region 2, and Framework Region 3 on Human VH3-Encoded Igs. *J. Immunol.* **157,** 2982–2988 (1996).

150. Palmqvist, N., Patti, J. M., Tarkowski, A. & Josefsson, E. Expression of Staphylococcal Clumping Factor A Impedes Macrophage Phagocytosis. *Microbes Infect.* **6,** 188–195 (2004).

151. Roghmann, M. *et al.* Epidemiology of Capsular and Surface Polysaccharide in *Staphylococcus aureus* Infections Complicated by Bacteraemia. *J. Hosp. Infect.* **59,** 27–32 (2005).

152. Nilsson, I. M., Lee, J. C., Bremell, T., Rydén, C. & Tarkowski, A. The Role of Staphylococcal Polysaccharide Microcapsule Expression in Septicemia and Septic Arthritis. *Infect. Immun.* **65,** 4216–4221 (1997).

153. Thakker, M., Park, J. S., Carey, V. & Lee, J. C. *Staphylococcus aureus* Serotype 5 Capsular Polysaccharide is Antiphagocytic and Enhances Bacterial Virulence in a Murine Bacteremia Model. *Infect. Immun.* **66,** 5183–5189 (1998).

154. Luong, T. T. & Lee, C. Y. Overproduction of Type 8 Capsular Polysaccharide Augments *Staphylococcus aureus* Virulence. *Infect. Immun.* **70,** 3389–3395 (2002).

155. Lee, J. C., Park, J.-S., Shepherd, S. E. & Fattom, A. Protective Efficacy of Antibodies to the *Staphylococcus aureus* Capsular Polysaccharide in a Rat Model of Endocarditis. *Infect. Immun.* **65,** 4146–4151 (1997).

156. Peschel, A. *et al.* Inactivation of the *dlt* operon in *Staphylococcus aureus* Confers Sensitivity to Defensins, Protegrins, and Other Antimicrobial Peptides. *J. Biol. Chem.* **274,** 8405–8410 (1999).

157. Peschel, A. *et al. Staphylococcus aureus* Resistance to Human Defensins and Evasion of Neutrophil Killing via the Novel Virulence Factor Mprf Is Based on Modification of Membrane Lipids with L-Lysine. *J. Exp. Med.* **193,** 1067–1076 (2001).

158. Jin, T. *et al. Staphylococcus aureus* Resists Human Defensins by Production of Staphylokinase, a Novel Bacterial Evasion Mechanism. *J. Immunol.* **172,** 1169–1176 (2004).

159. Sieprawska-Lupa, M. *et al.* Degradation of Human Antimicrobial Peptide LL-37 by *Staphylococcus aureus*-Derived Proteinases. *Antimicrob. Agents Chemother.* **48,** 4673–4679 (2004).

160. Gresham, H. D. *et al.* Survival of *Staphylococcus aureus* Inside Neutrophils Contributes to Infection. *J. Immunol.* **164,** 3713–3722 (2000).

161. Cheung, A. L., Koomey, J. M., Butler, C. A., Projan, S. J. & Fischetti, V. A. Regulation of Exoprotein Expression in *Staphylococcus aureus* by a locus (*sar*) Distinct from *agr*. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 6462–6466 (1992).

162. Liu, G. Y. *et al. Staphylococcus aureus* Golden Pigment Impairs Neutrophil Killing and Promotes Virulence Through Its Antioxidant Activity. *J. Exp. Med.* **202,** 209–215 (2005).

163. Clauditz, A., Resch, A., Wieland, K. P., Peschel, A. & Götz, F. Staphyloxanthin Plays a Role in the Fitness of *Staphylococcus aureus* and Its Ability to Cope with Oxidative Stress. *Infect. Immun.* **74,** 4950–4953 (2006).

164. Horsburgh, M. J. *et al.* MntR Modulates Expression of the PerR Regulon and Superoxide Resistance in *Staphylococcus aureus* Through Control of Manganese Uptake. *Mol. Microbiol.* **44,** 1269–1286 (2002).

165. Karavolos, M. H., Horsburgh, M. J., Ingham, E. & Foster, S. J. Role and Regulation of the Superoxide Dismutases of *Staphylococcus aureus*. *Microbiology* **149,** 2749–2758 (2003).

166. Singh, V. K. & Moskovitz, J. Multiple Methionine Sulfoxide Reductase Genes in *Staphylococcus aureus*: Expression of Activity and Roles in Tolerance of Oxidative Stress. *Microbiology* **149,** 2739–2747 (2003).

167. Weissbach, H. *et al.* Peptide Methionine Sulfoxide Reductase: Structure, Mechanism of Action, and Biological Function. *Arch. Biochem. Biophys.* **397,** 172–178 (2002).

168. Gross, M., Cramton, S. E., Götz, F. & Peschel, A. Key Role of Teichoic Acid Net Charge in *Staphylococcus aureus* Colonization of Artificial Surfaces Key Role of Teichoic Acid Net Charge in *Staphylococcus aureus* Colonization of Artificial Surfaces. *Infect. Immun.* **69,** 3423–3426 (2001).

169. Kennedy, C. A. & O'Gara, J. P. Contribution of Culture Media and Chemical Properties of Polystyrene Tissue Culture Plates to Biofilm Development by *Staphylococcus aureus*. *J. Med. Microbiol.* **53,** 1171–1173 (2004).

170. Fey, P. D. *et al.* A Genetic Resource for Rapid and Comprehensive Phenotype Screening of Nonessential *Staphylococcus aureus* Genes. *MBio* **4,** e00537-12-e00537-12 (2013).

171. Speziale, P., Pietrocola, G., Foster, T. J. & Geoghegan, J. A. Protein-Based Biofilm Matrices in Staphylococci. *Front. Cell. Infect. Microbiol.* **4,** 1–10 (2014).

172. Moormeier, D. E., Bose, J. L., Horswill, A. R. & Bayles, K. W. Temporal and Stochastic Control of *Staphylococcus aureus* Biofilm Development. *MBio* **5,** e01341-14 (2014).

173. Foulston, L., Elsholz, A. K. W., DeFrancesco, A. S. & Losick, R. The Extracellular Matrix of *Staphylococcus aureus* Biofilms Comprises Cytoplasmic Proteins That Associate with the Cell Surface in Response to Decreasing pH. *MBio* **5,** e01667-14 (2014).

174. Rice, K. C. *et al.* The *cidA* Murein Hydrolase Regulator Contributes to DNA Release and Biofilm Development in *Staphylococcus aureus*. *Proc. Natl. Acad. Sci.* **104,** 8113–8118 (2007).

175. Dengler, V., Foulston, L., DeFrancesco, A. S. & Losick, R. An Electrostatic Net Model for the Role of Extracellular DNA in Biofilm Formation by *Staphylococcus aureus*. *J. Bacteriol.* **197,** 3779–3787 (2015).

176. Mackey-Lawrence, N. M., Potter, D. E., Cerca, N. & Jefferson, K. K. *Staphylococcus aureus* Immunodominant Surface Antigen B is a Cell-Surface Associated Nucleic Acid Binding Protein. *BMC Microbiol.* **9,** (2009).

177. Huseby, M. J. *et al.* Beta Toxin Catalyzes Formation of Nucleoprotein Matrix in Staphylococcal Biofilms. *Proc. Natl. Acad. Sci.* **107,** 14407–14412 (2010).

178. Schwartz, K., Ganesan, M., Payne, D. E., Solomon, M. J. & Boles, B. R. Extracellular DNA Facilitates the Formation of Functional Amyloids in *Staphylococcus aureus* Biofilms. *Mol. Microbiol.* **99,** 123–134 (2016).

179. Cramton, S. E., Gerke, C., Schnell, N. F., Nichols, W. W. & Götz, F. The Intercellular Adhesion (*ica*) Locus is Present in *Staphylococcus aureus* and is Required for Biofilm Formation. *Infect. Immun.* **67,** 5427–5433 (1999).

180. Zapotoczna, M., McCarthy, H., Rudkin, J. K., O'Gara, J. P. & O'Neill, E. An Essential Role for Coagulase in *Staphylococcus aureus* Biofilm Development Reveals New Therapeutic Possibilities for Device-Related Infections. *J. Infect. Dis.* **212,** 1883–1893 (2015).

181. Peschel, A. & Otto, M. Phenol-Soluble Modulins and Staphylococcal Infection. *Nat. Rev. Microbiol.* **11,** 667–673 (2013).

182. Periasamy, S. *et al.* How *Staphylococcus aureus* Biofilms Develop Their Characteristic Structure. *Proc. Natl. Acad. Sci.* **109,** 1281–1286 (2012).

183. Moormeier, D. E. *et al.* Use of Microfluidic Technology to Analyze Gene Expression During *Staphylococcus aureus* Biofilm Formation Reveals Distinct Physiological Niches. *Appl. Environ. Microbiol.* **79,** 3413–3424 (2013).

184. Mann, E. E. *et al.* Modulation of eDNA Release and Degradation Affects *Staphylococcus aureus* Biofilm Maturation. *PLoS One* **4,** e5822 (2009).

185. Boles, B. R. & Horswill, A. R. *agr*-Mediated Dispersal of *Staphylococcus aureus* Biofilms. *PLoS Pathog.* **4,** e1000052 (2008).

186. Recsei, P. *et al.* Regulation of Exoprotein Gene Expression in *Staphylococcus aureus* by *agr*. *Mol. Gen. Genet.* **202,** 58–61 (1986).

187. Vuong, C., Saenz, H. L., Götz, F. & Otto, M. Impact of the *agr* Quorum-Sensing System on Adherence to Polystyrene in *Staphylococcus aureus*. *J. Infect. Dis.* **182,** 1688–1693 (2000).

188. Yarwood, J. M., Bartels, D. J., Volper, E. M. & Greenberg, E. P. Quorum Sensing in *Staphylococcus aureus* Biofilms. *J. Bacteriol.* **186,** 1838–1850 (2004).

189. Kim, M. K., Ingremeau, F., Zhao, A., Bassler, B. L. & Stone, H. A. Local and Global Consequences of Flow on Bacterial Quorum Sensing. *Nat. Microbiol.* **1,** 1–5 (2016).

190. Dinges, M. M., Orwin, P. M. & Schlievert, P. M. Exotoxins of *Staphylococcus aureus*. *Clin. Microbiol. Rev.* **13,** 16–34 (2000).

191. Laabei, M. *et al.* Evolutionary Trade-Offs Underlie the Multi-faceted Virulence of *Staphylococcus aureus*. *PLoS Biol.* **13,** e1002229 (2015).

192. Rudkin, J. K., McLoughlin, R. M., Preston, A. & Massey, R. C. Bacterial Toxins: Offensive, Defensive, or Something Else Altogether? *PLoS Pathog.* **13,** e1006452 (2017).

193. Das, S. *et al.* Natural Mutations in a *Staphylococcus aureus* Virulence Regulator Attenuate Cytotoxicity but Permit Bacteremia and Abscess Formation. *Proc. Natl. Acad. Sci.* **113,** E3101–E3110 (2016).

194. Joo, H. S., Cheung, G. Y. C. & Otto, M. Antimicrobial Activity of Community-Associated Methicillin-Resistant *Staphylococcus aureus* is Caused by Phenol-Soluble Modulin Derivatives. *J. Biol. Chem.* **286,** 8933–8940 (2011).

195. Alonzo III, F. & Torres, V. J. The Bicomponent Pore-Forming Leucocidins of *Staphylococcus aureus*. *Microbiol. Mol. Biol. Rev.* **78,** 199–230 (2014).

196. Song, L. *et al.* Structure of Staphylococcal α-Hemolysin, a Heptameric Transmembrane Pore. *Science (80-.).* **274,** 1859–1865 (1996).

197. Bhakdil, S. & Tranum-Jensen, J. Alpha-Toxin of *Staphylococcus aureus*. *Microbiol. Mol. Biol. Rev* **55,** 733–751 (1991).

198. Wilke, G. A. & Burbeck Wardenburg, J. Role of a Disintegrin and Metalloprotease 10 in *Staphylococcus aureus* α-Hemolysin-Mediated Cellular Injury. *Proc. Natl. Acad. Sci.* **107,** 13473–13478 (2010).

199. Seilie, E. S. & Bubeck Wardenburg, J. *Staphylococcus aureus* Pore-Forming Toxins: the Interface of Pathogen and Host Complexity. *Semin. Cell Dev. Biol.* **71,** 101–116 (2017).

200. Peacock, S. J. *et al.* Virulent Combinations of Adhesin and Toxin Genes in Natural Populations of *Staphylococcus aureus*. *Infect. Immun.* **70,** 4987–4996 (2002).

201. Shukla, S. K. *et al.* Virulence Genes and Genotypic Associations in Nasal Carriage, Community-Associated Methicillin-Susceptible and Methicillin-Resistant USA400 *Staphylococcus aureus* Isolates. *J. Clin. Microbiol.* **48,** 3582–3592 (2010).

202. Alonzo III, F. *et al. Staphylococcus aureus* Leucocidin ED Contributes to Systemic Infection by Targeting Neutrophils and Promoting Bacterial Growth *in vivo*. *Mol. Microbiol.* **83,** 423–435 (2012).

203. Gillet, Y. *et al.* Association Between *Staphylococcus aureus* Strains Carrying Gene for Panton-Valentine Leukocidin and Highly Lethal Necrotising Pneumonia in Young Immunocompetent Patients. *Lancet* **359,** 753–759 (2002).

204. Shallcross, L. J., Fragaszy, E., Johnson, A. M. & Hayward, A. C. The Role of the Panton-Valentine Leucocidin Toxin in Staphylococcal disease: a Systematic Review and Meta-Analysis. *Lancet Infect. Dis.* **13,** 43–54 (2013).

205. Vandenesch, F., Lina, G. & Henry, T. *Staphylococcus aureus* Hemolysins, Bi-Component Leukocidins, and Cytolytic Peptides: a Redundant Arsenal of Membrane-Damaging Virulence Factors? *Front. Cell. Infect. Microbiol.* **2,** (2012).

206. Wang, R. *et al.* Identification of Novel Cytolytic Peptides as Key Virulence Determinants for Community-Associated MRSA. *Nat. Med.* **13,** 1510–1514 (2007).

207. Pader, V. *et al. Staphylococcus aureus* Inactivates Daptomycin by Releasing Membrane Phospholipids. *Nat. Microbiol.* **2,** 16194 (2016).

208. Joo, H.-S. *et al.* Mechanism of Gene Regulation by a *Staphylococcus aureus* Toxin. *MBio* **7,** e01579-16 (2016).

209. Davis, J. P., Chesney, P. J., Wand, P. J., LaVenture, M. & Team, T. I. and L. Toxic-Shock Syndrome - Epidemiologic Features, Recurrence, Risk Factors, and Prevention. *N. Engl. J. Med.* **303,** 1429–1435 (1980).

210. Chesney, P. J. Clinical Aspects and Spectrum of Illness of Toxic Shock Syndrome: Overview. *Rev. Infect. Dis.* **11,** S1–S7 (1989).

211. Schlievert, P. M., Shands, K. N., Dan, B. B., Schmid, G. P. & Nishimura, R. D. Identification and Characterization of an Exotoxin from *Staphylococcus aureus* Associated with Toxic-Shock Syndrome. *J. Infect. Dis.* **143,** 509–516 (1981).

212. Bohach, G. A., Fast, D. J., Nelson, R. D. & Schlievert, P. M. Staphylococcal and Streptococcal Pyrogenic Toxins Involved in Toxic Shock Syndrome and Related Illnesses. *Crit. Rev. Microbiol.* **17,** 251–272 (1990).

213. Kent, T. H. Staphylococcal Enterotoxin Gastroenteritis in Rhesus Monkeys. *Am. J. Pathol.* **48,** 387–407 (1966).

214. Otto, M. *Staphylococcus aureus* Toxins. *Curr. Opin. Microbiol.* **17,** 32–37 (2014).

215. Cassat, J. E. *et al.* A Secreted Bacterial Protease Tailors the *Staphylococcus aureus* Virulence Repertoire to Modulate Bone Remodeling During Osteomyelitis. *Cell Host Microbe* **13,** 759–772 (2013).

216. Jusko, M. *et al.* Staphylococcal Proteases Aid in Evasion of the Human Complement System. *J. Innate Immun.* **6,** 31–46 (2014).

217. Bukowski, M., Wladyka, B. & Dubin, G. Exfoliative Toxins of *Staphylococcus aureus*. *Toxins (Basel).* **2,** 1148–1165 (2010).

218. Brinkmann, V. & Zychlinsky, A. Neutrophil Extracellular Traps: is Immunity the Second Function of Chromatin? *J. Cell Biol.* **198,** 773–783 (2012).

219. Rosenstein, R. & Götz, F. Staphylococcal Lipases: Biochemical and Molecular Characterization. *Biochimie* **82,** 1005–1014 (2000).

220. Neumann, Y. *et al.* The Effect of Skin Fatty Acids on *Staphylococcus aureus*. *Arch. Microbiol.* **197,** 245–267 (2015).

221. Hu, C., Xiong, N., Zhang, Y., Rayner, S. & Chen, S. Functional Characterization of Lipase in the Pathogenesis of *Staphylococcus aureus*. *Biochem. Biophys. Res. Commun.* **419,** 617–620 (2012).

222. Patil, K. J., Chopda, M. Z. & Mahajan, R. T. Lipase Biodiversity. *Indian J. Sci. Technol.* **4,** 971–982 (2011).

223. Bronner, S., Monteil, H. & Prévost, G. Regulation of Virulence Determinants in *Staphylococcus aureus*: Complexity and Applications. *FEMS Microbiol. Rev.* **28,** 183–200 (2004).

224. Laabei, M. & Massey, R. C. Using Functional Genomics to Decipher the Complexity of Microbial Pathogenicity. *Curr. Genet.* **62,** 523–525 (2016).

225. Thoendel, M., Kavanaugh, J. S., Flack, C. E. & Horswill, A. R. Peptide Signaling in the Staphylococci. *Chem. Rev.* **111,** 117–151 (2011).

226. Abdelnour, A., Arvidson, S., Bremell, T., Ryden, C. & Tarkowski, A. The Accessory Gene Regulator (*agr*) Controls *Staphylococcus aureus* Virulence in a Murine Arthritis Model. *Infect. Immun.* **61,** 3879–3885 (1993).

227. Traber, K. E. *et al. agr* Function in Clinical *Staphylococcus aureus* Isolates. *Microbiology* **154,** 2265–2274 (2008).

228. Cheung, G. Y. C., Wang, R., Khan, B. A., Sturdevant, D. E. & Otto, M. Role of the Accessory Gene Regulator *agr* in Community-Associated Methicillin-Resistant *Staphylococcus aureus* Pathogenesis. *Infect. Immun.* **79,** 1927–1935 (2011).

229. McNamara, P. J. & Bayer, A. S. A *rot* Mutation Restores Parental Virulence to an *agr*-Null *Staphylococcus aureus* Strain in a Rabbit Model of Endocarditis. *Infect. Immun.* **73,** 3806–3809 (2005).

230. Painter, K. L., Krishna, A., Wigneshweraraj, S. & Edwards, A. M. What Role Does the Quorum-Sensing Accessory Gene Regulator System Play During *Staphylococcus aureus* Bacteremia? *Trends Microbiol.* **22,** 676–685 (2014).

231. Shopsin, B. *et al.* Prevalence of agr Dysfunction among Colonizing *Staphylococcus aureus* Strains. *J. Infect. Dis.* **198,** 1171–1174 (2008).

232. Peterson, M. M. *et al.* Apolipoprotein B is an Innate Barrier against Invasive *Staphylococcus aureus* Infection. *Cell Host Microbe* **4,** 555–566 (2008).

233. Malachowa, N. *et al.* Global Changes in *Staphylococcus aureus* Gene Expression in Human Blood. *PLoS One* **6,** e18617 (2011).

234. Surewaard, B. G. J. *et al.* Inactivation of Staphylococcal Phenol Soluble Modulins by Serum Lipoprotein Particles. *PLoS Pathog.* **8,** e1002606 (2012).

235. Rothfork, J. M., Dessus-Babus, S., Van Wamel, W. J. B., Cheung, A. L. & Gresham, H. D. Fibrinogen Depletion Attenuates *Staphyloccocus aureus* Infection by Preventing Density-Dependent Virulence Gene Up-Regulation. *J. Immunol.* **171,** 5389–5395 (2003).

236. James, E. H., Edwards, A. M. & Wigneshweraraj, S. Transcriptional Downregulation of *agr* Expression in *Staphylococcus aureus* During Growth in Human Serum Can Be Overcome by Constitutively Active Mutant Forms of the Sensor Kinase AgrC. *FEMS Microbiol. Lett.* **349,** 153–162 (2013).

237. Goerke, C., Fluckiger, U., Steinhuber, A., Zimmerli, W. & Wolz, C. Impact of the Regulatory Loci *agr*, *sarA* and *sae* of *Staphylococcus aureus* on the Induction of α-toxin During Device-Related Infection Resolved by Direct Quantitative Transcript analysis. *Mol. Microbiol.* **40,** 1439–1447 (2001).

238. Baroja, M. L. *et al.* The SaeRS Two-Component System is a Direct and Dominant Transcriptional Activator of Toxic Shock Syndrome Toxin 1 in *Staphylococcus aureus*. *J. Bacteriol.* **198,** 2732–2742 (2016).

239. Read, T. D. & Massey, R. C. Characterizing the Genetic Basis of Bacterial Phenotypes Using Genome-Wide Association Studies: a New Direction for Bacteriology. *Genome Med.* **6,** 109 (2014).

240. Slatkin, M. Linkage Disequilibrium - Understanding the Evolutionary Past and Mappping the Medical Future. *Nat. Rev. Genet.* **9,** 477–485 (2008).

241. Cha, R. S., Zarbl, H., Keohavong, P. & Thilly, W. G. Mismatch Amplification Mutation Assay (MAMA): Application to the c-H-ras Gene. *Genome Res.* **2,** 14–20 (1992).

242. Sowell, J., Pollard, L. & Wood, T. Quantification of Branched-Chain Amino Acids in Blood Spots and Plasma by Liquid Chromatography Tandem Mass Spectrometry for the Diagnosis of Maple Syrup Urine Disease. *J. Sep. Sci.* **34,** 631–639 (2011).

243. Corrigan, R. M. & Foster, T. J. An Improved Tetracycline-Inducible Expression Vector for *Staphylococcus aureus*. *Plasmid* **61,** 126–129 (2009).

244. Reiss, S. *et al.* Global Analysis of the *Staphylococcus aureus* Response to Mupirocin. *Antimicrob. Agents Chemother.* **56,** 787–804 (2012).

245. Geiger, T. *et al.* The Stringent Response of *Staphylococcus aureus* and its Impact on Survival After Phagocytosis Through the Induction of Intracellular PSMs Expression. *PLoS Pathog.* **8,** e1003016 (2012).

246. Laabei, M. *et al.* Predicting the Virulence of MRSA from Its Genome Sequence. *Genome Res.* **24,** 839–849 (2014).

247. Kennedy, A. D. *et al.* Complete Nucleotide Sequence Analysis of Plasmids in Strains of *Staphylococcus aureus* Clone USA300 Reveals a High Level of Identity among Isolates with Closely Related Core Genome Sequences. *J. Clin. Microbiol.* **48,** 4504–4511 (2010).

248. Diep, B. A. *et al.* Complete Genome Sequence of USA300, an Epidemic Clone of Community-Acquired Meticillin-Resistant *Staphylococcus aureus*. *Lancet* **367,** 731–9 (2006).

249. Hurdle, J. G., O'Neil, A. J. & Chopra, I. The Isoleucyl-tRNA Synthetase Mutation V588F Conferring Mupirocin Resistance in Glycopeptide-Intermediate *Staphylococcus aureus* is Not Associated with a Significant Fitness Burden. *J. Antimicrob. Chemother.* **53,** 102–104 (2003).

250. Moura de Sousa, J., Balbontín, R., Durão, P. & Gordo, I. Multidrug-Resistant Bacteria Compensate for the Epistasis Between Resistances. *PLOS Biol.* **15,** e2001741 (2017).

251. Antonio, M., McFerran, N. & Pallen, M. J. Mutations Affecting the Rossman fold of Isoleucyl-tRNA Synthetase are Correlated with Low-Level Mupirocin Resistance in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **46,** 438–442 (2002).

252. Mooney, C. Z. *Monte Carlo Simulations*. (SAGE Publications, 1997).

253. Holden, M. T. G. *et al.* Genome Sequence of a Recently Emerged, Highly Transmissible, Multi-Antibiotic- and Antiseptic-Resistant Variant of Methicillin-Resistant *Staphylococcus aureus*, Sequence Type 239 (TW). *J. Bacteriol.* **192,** 888–892 (2010).

254. Takigawa, H., Nakagawa, H., Kuzukawa, M. & Imokawa, G. Deficient Production of Hexadecenoic Acid in the Skin Is Associated in Part with the Vulnerability of Atopic Dermatitis Patients to Colonization by *Staphylococcus aureus*. *Dermatology* **211,** 240–248 (2005).

255. Shryock, T. R. & Kapral, F. A. The Production of Bactericidal Fatty Acids from Glycerides in Staphylococcal Abscesses. *J. Med. Microbiol.* **36,** 288–292 (1992).

256. White, M. J., Boyd, J. M., Horswill, a. R. & Nauseef, W. M. Phosphatidylinositol-Specific Phospholipase C Contributes to Survival of *Staphylococcus aureus* USA300 in Human Blood and Neutrophils. *Infect. Immun.* **82,** 1559–1571 (2014).

257. Rollof, J., Hedström, S. Å. & Nilsson-Ehle, P. Lipolytic Activity of *Staphylococcus aureus* Strains from Disseminated and Localized Infections. *Acta Pathol. Microbiol. Scand. Ser. B Microbiol.* **95B,** 109–113 (1987).

258. Rollof, J., Braconier, J. H., Söderström, C. & Nilsson-Ehle, P. Interference of *Staphylococcus aureus* Lipase with Human Granulocyte Function. *Eur. J. Clin. Microbiol. Infect. Dis.* **7,** 505–510 (1988).

259. Blevins, J. S., Beenken, K. E., Elasri, M. O., Hurlburt, B. K. & Smeltzer, M. S. Strain-Dependent Differences in the Regulatory Roles of *sarA* and *agr* in *Staphylococcus aureus*. *Infect. Immun.* **70,** 470–480 (2002).

260. Ben Bacha, A., Moubayed, N. M. S. & Al-Assaf, A. An Organic Solvent-Stable Lipase from a Newly Isolated *Staphylococcus aureus* ALA1 Strain with Potential for Use as an Industrial Biocatalyst. *Biotechnol. Appl. Biochem.* **63,** 378–390 (2016).

261. Gupta, N., Rathi, P. & Gupta, R. Simplified *para*-Nitrophenyl Palmitate Assay for Lipases and Esterases. *Anal. Biochem.* **311,** 98–99 (2002).

262. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).

263. Burlak, C. *et al.* Global Analysis of Community-Associated Methicillin-Resistant *Staphylococcus aureus* Exoproteins Reveals Molecules Produced *in vitro* and During Infection. *Cell. Microbiol.* **9,** 1172–1190 (2007).

264. O'Neill, E. *et al.* A Novel *Staphylococcus aureus* Biofilm Phenotype Mediated by the Fibronectin-Binding Proteins, FnBPA and FnBPB. *J. Bacteriol.* **190,** 3835–3850 (2008).

265. Silver, L. L. Fosfomycin: Mechanism and Resistance. *Cold Spring Harb. Perspect. Med.* **7,** a025262 (2017).

266. Hazard, A. L. & Senior, A. E. Defective Energy Coupling in δ-Subunit Mutants of *Escherichia coli* F1F0-ATP Synthase. *J. Biol. Chem.* **269,** 427–432 (1994).

267. Saibil, H. R. Chaperone Machines in Action. *Curr. Opin. Struct. Biol.* **18,** 35–42 (2008).

268. Sheppard, S. K. *et al.* Genome-Wide Association Study Identifies Vitamin B5 Biosynthesis as a Host Specificity factor in *Campylobacter*. *Proc. Natl. Acad. Sci.* **110,** 11923–11927 (2013).

269. Farhat, M. R. *et al.* Genomic Analysis Identifies Targets of Convergent Positive Selection in Drug-Resistant *Mycobacterium tuberculosis*. *Nat. Genet.* **45,** 1183–1189 (2013).